

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:46:32 ; Search time 504 Seconds  
(without alignments)

6024.609 Million cell updates/sec

Title: US-10-761-033-1\_COPY\_1\_1316

Perfect score: 1316

Sequence: 1 atgacaggagcaaggagaa.....gtgataaaatcgattgggg 1316

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 115364844 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications NA New:  
1: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214.2	16.3	2829	7	US-10-505-263-88
2	203.2	15.4	3635	7	US-10-505-263-80
3	184.6	14.0	3200	7	US-10-505-263-86
4	178.2	13.5	197096	12	US-11-121-086-107
5	177.7	13.4	2640	7	US-10-505-263-82
6	176.6	13.4	3367	7	US-10-505-263-11
7	172	13.1	4815	7	US-10-537-002-56
8	168.8	12.8	2654	7	US-10-505-263-1
9	168.8	12.8	2748	7	US-10-505-263-3
10	166.4	12.6	2279	7	US-10-505-263-90
11	164.2	12.5	2612	7	US-10-505-263-5
12	164.2	12.5	2750	7	US-10-505-263-7
13	157.8	12.0	2457	7	US-10-505-263-84
14	109	8.3	2871	7	US-10-505-263-9
15	48.6	3.7	2290	9	US-11-096-568A-33954
16	44.8	3.4	2298	9	US-11-096-568A-31695
17	44	3.3	10490	12	US-11-011-332A-87
18	43.4	3.3	10490	12	US-11-011-332A-101
19	43.2	3.3	610	6	US-09-925-065A-422255
20	43.2	3.3	610	6	US-09-925-065A-422256

21	43	3.3	1452	8	US-10-793-626-3131	Sequence 3131, Ap
22	43	3.3	3926	8	US-10-793-626-4300	Sequence 4300, Ap
23	43	3.3	4189	8	US-10-793-626-3898	Sequence 3898, Ap
24	42.8	3.3	610	6	US-09-925-065A-422254	Sequence 422254,
25	41.6	3.2	568	6	US-09-925-065A-264855	Sequence 264855,
26	40.4	3.1	2417	12	US-11-010-239-47	Sequence 47, Appl
27	39.8	3.0	1004	8	US-10-793-626-4406	Sequence 4406, Ap
28	39.6	3.0	503	6	US-09-925-065A-558945	Sequence 558945,
29	39.6	3.0	548	6	US-09-925-065A-558944	Sequence 558944,
30	39.6	3.0	888	7	US-10-932-182A-3583	Sequence 3583, Ap
31	39.6	3.0	888	7	US-10-932-182A-3583	Sequence 3583, Ap
32	39.6	3.0	1695	12	US-11-098-686-9652	Sequence 9652, Ap
33	39.6	3.0	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
34	39	3.0	643	6	US-09-925-065A-769517	Sequence 769517,
35	38.8	2.9	2685	7	US-10-932-182A-1726	Sequence 1726, Ap
36	38.8	2.9	2685	7	US-10-932-182A-1726	Sequence 1726, Ap
37	38.2	2.9	640	6	US-09-925-065A-623209	Sequence 623209,
38	38	2.9	770	7	US-10-349-331-225	Sequence 225, App
39	37.8	2.9	890	12	US-11-136-527-1445	Sequence 1445, Ap
40	37.8	2.9	890	12	US-11-136-527-1445	Sequence 5541, Ap
41	37.6	2.9	148220	12	US-11-121-086-90	Sequence 90, Appl
42	37.6	2.9	179597	12	US-11-121-086-91	Sequence 91, Appl
43	37.2	2.8	524	6	US-09-925-065A-405101	Sequence 405101,
44	37	2.8	148935	8	US-10-995-561-13308	Sequence 13308, A
45	37	2.8	1082144	12	US-11-117-187-211	Sequence 211, App

#### ALIGNMENTS

#### RESULT 1

US-10-505-263-88  
; Sequence 88, Application US/10505263  
; Publication No. US20060014940A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Case Western Reserve University  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Mount, David B  
; APPLICANT: Romero, Michael  
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, and SLC26A2  
; FILE REFERENCE: 1342/50/2 PCT/US  
; CURRENT APPLICATION NUMBER: US/10/505,263  
; PRIOR FILING DATE: 2004-08-20  
; PRIOR APPLICATION NUMBER: US 60/360,275  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: PCT/US03/06469  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88  
; LENGTH: 2829  
; TYPE: DNA  
; ORGANISM: Xenopus laevis  
; US-10-505-263-88

Query Match 16.3%; Score 214.2; DB 7; Length 2829;

Best Local Similarity 49.1%; Pred. No. 2.8e-46;  
Matches 598; Conservative 0; Mismatches 613; Indels 6; Gaps 1;

QY	95	TGCCCATTTTGGATTGGGCACCACTTACATCTGAAAGAAAACCTTGTTCAGACACTG	154
DB	521	TTCCAAATCTTGGTTGGTTACCGGTATCCAGTAAGGAATGGCTGCTCGAGACATAG	580
QY	155	TGCTGGGATAATGTGGCACTTCAACAGGTGACCAAGGATGGCCTTTGCTTCTCT	214
DB	581	TATCTGGTCTTAGTGTGGCATCATTCAGCTCCCAAGGTTTACTTATGCTTTACTGG	640
QY	215	CATCTGTCACCCAGTGTTCGTTTATATGGTCTCTCTGTTCCGCCCAATATTTATGCCA	274
DB	641	CTGAGATTCCACCGGTGTTGGATTGTTATCTCTATTTTCCCTACTTCTCTATGCCA	700

```
QY 275 TATTTGGAAATGGGACATCATGTTGGCCACAGGACACTTTGGCTTGACATCTCTTAATATACAG 334
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 701 TATTTGGGACTTCAAGACATATTTCCCTCGGAACCTTTGCTGTGATTTCTGTGATGGTTG 760
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 335 CCAACGCCGTGGAAACGGATTTCCCTCAGAACATGCAGAACTCCACACACAGAGTACA 394
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 761 GTAGTGTGACTGAATCACTGTGGTCC-----GTGAGAAATCTACAGGTTACTCGGAATG 814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 395 CAAGCGTGTGGCTTATCCGACTTTGAAATGCAAAAGATCCACGTTGCTGCGACAGTTT 454
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 AATCAGTGAATGATATAGCAGCAGAGACAATGACAGAGTTGAGTGGCATCGGCTTGA 874
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 455 CTTCTTGGGAGGTGATTCAGTGGCCATGTTTGTCTGCAACTGGGCGAGTGCACAT 514
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 CATCTAGTTGGACTTTTTCAGATTATGTTGGGTCTGTGCAAGTTGGCTTTGTTGTA 934
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 515 TTGTGTCACAGAGCTGTGATCAGCGCAATGACAACTGGGGCTGCCACCCATGTGTGA 574
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 CCTACCTCTCTGAACCTTGATCCGAGGCTACCCAGTGTGCTGCAATCCATGTCACTG 994
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 575 CTTCAAGTCAATATCTCTTGGGAATGAAATGCCATATATATCCGGACCACTTGGAT 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 995 TGTCTCAGATGAAGTGTCTTAGCGGTGCAGATCAGCAAGAGTCAATCATGTGCTC 1054
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 TCTTTTATATTATGATATGTTTGTGAAATCATCAAGTCTGTGCGACTGGAAGCATTGC 694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1055 TAAATTTATGCAATCTGTGCAAGTCTCTGTGCAAGTCTCTGAGCAAAATATGCTTCCCTTC 1114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 TTTTATCCTTCTGAGCAATGTGGTCTGTCTTGTGTTTAAAGAGCTGAATGAACAGTTTA 754
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1115 TAAATGGGTGCACTCCCATCACTGTCTGTCTGTGTTTCTGTGTTTAAAGTTCTTAAACGACAAGTACA 1174
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 AAAGGAAATTAAGTGTCTCTCTGTAGATTTAGTTTTCATATTGCTGCATCATTTG 814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1175 GCTCAAGATCGTATGCCCAATACCAATCGAATCATCACTCACTCACTGATAGCTGGA 1234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 CTTGTTATTTGCAACCAATATGAAACACATATGGAATAGAAATGATGTTGTCATATCCAC 874
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1235 TATCATATGTTGCCAGTCTCCACCAAGTCTATGAGTGGACATTTGTAGGGAGATGCCAA 1294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 AAGNATTCCTCACTAGAGCTCCCCGATGAACATCTCTCTGCGGTGATCACTGAG 934
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1295 CTGGATGAAGCTCCAATGTTACCAACACANAATATTTTTCGAAGATGGTTCGGCAATG 1354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 CTTTCGGAGTGGCACTTGTAGGCTATGTGGCTCACTGGCTCTTGTCTCAAGGATCTGCCA 994
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1355 CTTTGGCATCGCTGTTGTGTCTATGTCTTTACATCTCCTGTGCTAAGATGTTGGTG 1414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 995 AAAAAATCAAATATTCATATGATGACAAACAGGAATTTTGGCCCATGGCTCAGCAATA 1054
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1415 TTAAACATGGATACAAATATGACAGCAATCAGGAACGTGATCGCTTGGGGCTGAGTAAT 1474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1055 TAGTTTCTTCAATTTTCTTCTGATACCAAGTGTCTGCTCCATGGGAAGAGCGCTGCC 1114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1475 CAATAGGAAGCTTCTTTCAGTGTCTTCAATTTGGGACAGCCATGTCAAGAAATGCTGGTGC 1534
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1115 TGTACAGCACAGGACGGAAGACACAGGTGGCTGTCTTAATATCTTTCATTTTCTCTCTTA 1174
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1535 AAGAAAGCAAGAGGGGCAAGTCAAGTAGCAGTGTGTCTCTTCTTGTGTCACTCTTA 1594
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1175 TAGTCATCTATGCAATAGGACCTTTTGTCTTTATCTGGCTGCCCATGTGTCTCTTCAAGCA 1234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1595 TTATCATTTAAAGCTGGGAGCTGTTTGAACCCCTTACCAAAAGCCATCTAGCTGCAG 1654
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1235 TTATGTTGGGACTGAGGGAATGCTAATACAGTTCCGAGATTTAAAAAATATTGGA 1294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1655 TTGTTGTAGTCAATCTCAAGGTTATTTACAGAGCTTTTACTGATGTGCCAATGTTATGGA 1714
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1295 ATGTGGATAAAATCGAT 1311
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1715 GATCCAACAAAATTTGAT 1731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 2

```
US-10-505-263-80
; Sequence 80, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; TITLE OF INVENTION: ANION EXCHANGERS
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 80
; LENGTH: 3635
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-505-263-80
```

Query Match 15.4%; Score 203.2; DB 7; Length 3635;

Best Local Similarity 48.5%; Pred. No. 2.6e-43; Indels 9; Gaps 1;

Matches 596; Conservative 0; Mismatches 623;

```
QY 94 CTGCCCCATTTGGATTGGGCACCACTTACAACTCGAAGAAACTTGTCTTCAGACACT 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 CTCCTATTCTGAATTTGTTACCAATACAGATGGAAGAATGTTTTTAAGTGACCTG 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 GTGCTGGGATAATGTTGGCAGTTCAACAGGTGACCAAGGATTTGGCTTGTGTTCTC 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 ATCTCTGGAGTCAGCACAGGCTCTTGTGTGTACCTTTGCAAGGCTTGGCATTTGCTT 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 214 TCATCTGTGCACCCAGTGTGTTGTTATATGGGTCTCTGTTTCTGCCATAATTTATGCC 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GCTGAGTCCCTGTTGGATATGGCTCTACTCTTCTTTTCTCTATCTTGTGACTTATTTT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 ATATTTGGAATGGGACATCATGTTGGCCACAGGCACTTTGCTTTGACATCTTAAATCA 333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 TTCTGGGAACATCAAAACACATTTCTGTGGACCAATTCACAGTGGTCAGTTTGTATG 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 GCCAACGCCCTGGAAACGATTTGCTCCCTCAGAACATGACAGATCT-----CACACA 384
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 GGATCAGTTCTGTCTATTTGCTCCAGATGAAAAATTTGCAATATTAGGCAACTCTACA 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 385 CAGAGTAAACAACAGCGTCTGGGCTTATCCGACTTTGAAATGCAAGATCCAGTTGCT 444
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GGAAGTGAACAAAACAATAATAGACACAGTAGAAGATGCGAGCGAGTTGCCGTTTCA 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 445 GCAGCAGTTCTCTTCTTGGGAGGTGATTCAGGTGGCCATGTTTGTGCTGCAACTGGGC 504
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GGAACACTCAGCTTCTAATTGGNAATTATACAGCTTGCACTTGGCGTATTTCAAATTTGC 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 505 AGTGCCACATTTGTGTGTCACAGAGCTGTGATCAGCGCAATGACAACTGGGGCTGCAACC 564
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 TTTATTATCAGATATTGTCAGATCCATTAGTTGGTGTGTTTACTACAGCTGCAAGCTTC 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 565 CATGTGTGACTTTCACAAGTCAAAATCTCTTGGGAATGAAATGCCATATATATATCGGA 624
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 CAAGTGTGTTGTTTTCACAATTAATAAATAGATATCTTTACCAACATTTGCAACGGA 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 625 CCACCTTGGATCTTTTATATTTTATGCAATGTTTGTGAAAAACATCAAGTCTGTGCGACTG 684
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 GTACTCTCATATATATACAAATAAATAGATATCTTTACCAACATTTGCAAAACCAACATT 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 685 GAAGCATTTGCTTTATTCCTTGTGAGCAATTTGGTCTCTGTTCTTGTGTTAAGAGCTGAAT 744
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:44:50 ; Search time 1196 Seconds  
(without alignments)  
9099.081 Million cell updates/sec

Title: US-10-761-033-1\_COPY\_1\_1316

Perfect score: 1316  
Sequence: 1 atgacaggagcaagaggaa.....gtgataaaatcgattgggg 1316

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316	100.0	1797	3 US-09-875-811-11	Sequence 11, Appl
2	1316	100.0	1797	7 US-10-761-033-11	Sequence 11, Appl
3	1316	100.0	1818	3 US-09-875-811-7	Sequence 7, Appl
4	1316	100.0	1818	7 US-10-761-033-7	Sequence 7, Appl
5	1316	100.0	1866	3 US-09-875-811-3	Sequence 3, Appl
6	1316	100.0	1866	7 US-10-761-033-3	Sequence 3, Appl
7	1316	100.0	1971	3 US-09-875-811-9	Sequence 9, Appl
8	1316	100.0	1971	7 US-10-761-033-9	Sequence 9, Appl
9	1316	100.0	1971	9 US-10-505-316-1	Sequence 1, Appl
10	1316	100.0	1992	3 US-09-875-811-5	Sequence 5, Appl
11	1316	100.0	1992	7 US-10-761-033-5	Sequence 5, Appl
12	1316	100.0	2040	3 US-09-875-811-1	Sequence 1, Appl
13	1316	100.0	2040	7 US-10-761-033-1	Sequence 1, Appl
14	1316	100.0	2345	5 US-10-222-009-12	Sequence 12, Appl
15	1316	100.0	2420	3 US-09-875-811-13	Sequence 13, Appl
16	1316	100.0	2420	7 US-10-761-033-13	Sequence 13, Appl
17	1314.4	99.9	1971	3 US-09-942-446-3	Sequence 3, Appl
18	1314.4	99.9	2631	3 US-09-814-353-20997	Sequence 20997, A
19	1314.4	99.9	2886	5 US-10-222-009-11	Sequence 11, Appl
20	1314.4	99.9	5267	3 US-09-942-446-1	Sequence 1, Appl
21	1166.2	88.6	2848	7 US-10-415-378-24	Sequence 24, Appl
22	1068	81.2	2289	9 US-10-505-316-3	Sequence 3, Appl
23	514.4	39.1	1296	9 US-10-450-763-6290	Sequence 6290, Ap

ALIGNMENTS

RESULT 1

; Sequence 11, Application US/09875811  
; Patent No. US20020032321A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. US20020032321A1 Human Transporter Proteins and Polynucleotic  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; CURRENT FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1797  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-875-811-11

Query Match	100.0%	Score 1316;	DB 3;	Length 1797;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1316;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGACAGAGCAAAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCCAG	60	
Db	1	ATGACAGAGCAAAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCCAG	60	
Qy	61	TGTGAAGACATTTATACAGTGTGTAGAGGGAGCTGCCCATTTTGGATTGGCACCACAT	120	
Db	61	TGTGAAGACATTTATACAGTGTGTAGAGGGAGCTGCCCATTTTGGATTGGCACCACAT	120	
Qy	121	TACAACTCTGAAGAAAGAACTTCTTCCAGACACTGTGTCTGGGTAATATTTGGCAGTTCAA	180	
Db	121	TACAACTCTGAAGAAAGAACTTCTTCCAGACACTGTGTCTGGGTAATATTTGGCAGTTCAA	180	
Qy	181	CAGGTGACCCCAAGATTTGGCCCTTTCTCTCATCTGTGACCCAGTGTGGTTTGA	240	
Db	181	CAGGTGACCCCAAGATTTGGCCCTTTCTCTCATCTGTGACCCAGTGTGGTTTGA	240	
Qy	241	TATGGTCTCTGTTTCTCTGCCATAATTTATGCCATATTTGGAATGGGACATCATGTGCC	300	
Db	241	TATGGTCTCTGTTTCTCTGCCATAATTTATGCCATATTTGGAATGGGACATCATGTGCC	300	

```
QY 301 ACAGGACACCTTTGCTTGACATCTTAATATCAGCAACCGCGTGGAAACGGATTGTCCCT 360
Db 301 ACAGGACACCTTTGCTTGACATCTTAATATCAGCAACCGCGTGGAAACGGATTGTCCCT 360
QY 361 CAGAACATGCAGATCTCACACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT 420
Db 361 CAGAACATGCAGATCTCACACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT 420
QY 421 GAAATGCAAGGATCCAGTGTGCTGCAGCAGTTTCTTCTTGGGAGGTGATTCAGGTG 480
Db 421 GAAATGCAAGGATCCAGTGTGCTGCAGCAGTTTCTTCTTGGGAGGTGATTCAGGTG 480
QY 481 GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCACAGAGCCTGTGATCAGC 540
Db 481 GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCACAGAGCCTGTGATCAGC 540
QY 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCAACAGTCAAAATATCTCTTGGG 600
Db 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCAACAGTCAAAATATCTCTTGGG 600
QY 601 ATGAAATGCCATATATCCGGACCACTTGGATCTTTTATATTTATGATGATGTTT 660
Db 601 ATGAAATGCCATATATCCGGACCACTTGGATCTTTTATATTTATGATGATGTTT 660
QY 661 GAAACATCAAGTCTGTGGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGGTC 720
Db 661 GAAACATCAAGTCTGTGGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGGTC 720
QY 721 CTGTGTTCTTTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGAGTTGTTCTTCT 780
Db 721 CTGTGTTCTTTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGAGTTGTTCTTCT 780
QY 781 GTAGATTTAGTTTGAATTTAGTGTGCATCATTTGCTTATGACCAATATGGAAC 840
Db 781 GTAGATTTAGTTTGAATTTAGTGTGCATCATTTGCTTATGACCAATATGGAAC 840
QY 841 ACATATGGATTAGAATAGTTGGTCATATTCACAAAGGAAATTCCTCACTAGAGCTCCC 900
Db 841 ACATATGGATTAGAATAGTTGGTCATATTCACAAAGGAAATTCCTCACTAGAGCTCCC 900
QY 901 CCGATGAACATCCTCTCTGCGGTGATCATAGAGCTTTGGAGTGGCATCTGTAGGCTAT 960
Db 901 CCGATGAACATCCTCTCTGCGGTGATCATAGAGCTTTGGAGTGGCATCTGTAGGCTAT 960
QY 961 GTGGCTCAGTGGCTTTGCTCAAGGATCTGCCAAAATTCAAATATTCATATGATGAC 1020
Db 961 GTGGCTCAGTGGCTTTGCTCAAGGATCTGCCAAAATTCAAATATTCATATGATGAC 1020
QY 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCATTTTCTTCTGCATA 1080
Db 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCATTTTCTTCTGCATA 1080
QY 1081 CCAAGTGTGCTGCCATGGGAAGGACGGCTGCTGTAAGCAACAGAGGAGGACACACAG 1140
Db 1081 CCAAGTGTGCTGCCATGGGAAGGACGGCTGCTGTAAGCAACAGAGGAGGACACACAG 1140
QY 1141 GTGGCTGCTTAATATCTTGCAATTTTCTGCTTATAGTCACTATGCAATAGGACCTTTG 1200
Db 1141 GTGGCTGCTTAATATCTTGCAATTTTCTGCTTATAGTCACTATGCAATAGGACCTTTG 1200
QY 1201 CTTTACTGGCTGCCCATGTGTCTTCCAAAGCATTTATGTTGGGACTGAAGGGAATG 1260
Db 1201 CTTTACTGGCTGCCCATGTGTCTTCCAAAGCATTTATGTTGGGACTGAAGGGAATG 1260
QY 1261 CTAATACAGTCCGAGATTTAAAAAATTTGGAAATGGAATGGAATAAATCGATGGGG 1316
Db 1261 CTAATACAGTCCGAGATTTAAAAAATTTGGAAATGGAATGGAATAAATCGATGGGG 1316
```

RESULT 2

US-10-761-033-11

; Sequence 11, Application US/10761033

; Publication No. US20040138442A1

```
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/10/761.033
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/875,811
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-761-033-11
```

```
Query Match 100.0%; Score 1316; DB 7; Length 1797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAGGAGCAAAAGAGAGAAAAGAAAGCATGCTTTGGAGCAAGATGCATACCCCCCAG 60
Db 1 ATGACAGGAGCAAAAGAGAGAAAAGAAAGCATGCTTTGGAGCAAGATGCATACCCCCCAG 60
QY 61 TGTGAAGACATTATACAGTGTGTAGAAAGGCGACTGCCCATTTTGGATTGGGCAACAT 120
Db 61 TGTGAAGACATTATACAGTGTGTAGAAAGGCGACTGCCCATTTTGGATTGGGCAACAT 120
QY 121 TACAATCTGAAAGAAAACCTTCTCCAGACACTGTGTCTGGGTAATGTTGCGAGTTCAA 180
Db 121 TACAATCTGAAAGAAAACCTTCTCCAGACACTGTGTCTGGGTAATGTTGCGAGTTCAA 180
QY 181 CAGGTGACCCCAAGGATTTGGCTTTGCTCTCATCTGTGCACCCAGTGTGGTTTA 240
Db 181 CAGGTGACCCCAAGGATTTGGCTTTGCTCTCATCTGTGCACCCAGTGTGGTTTA 240
QY 241 TATGGGTCTGTGTTCTTCTGCCATTAATTTATGCCATATTTTGGAAATGGGACATCATGTTGCC 300
Db 241 TATGGGTCTGTGTTCTTCTGCCATTAATTTATGCCATATTTTGGAAATGGGACATCATGTTGCC 300
QY 301 ACAGGCACTTTTGGCTTGTGACATCTTAATATCAGCAACCGCGTGGAAACGGATTGTCCCT 360
Db 301 ACAGGCACTTTTGGCTTGTGACATCTTAATATCAGCAACCGCGTGGAAACGGATTGTCCCT 360
QY 361 CAGAACATGCAGAAATCTCACCAACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT 420
Db 361 CAGAACATGCAGAAATCTCACCAACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT 420
QY 421 GAAATGCAAGGATCCAGTGTGCTGACGAGTTTCTTCTTGGGAGGTGATTCAGGTG 480
Db 421 GAAATGCAAGGATCCAGTGTGCTGACGAGTTTCTTCTTGGGAGGTGATTCAGGTG 480
QY 481 GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCACAGAGCCTGTGATCAGC 540
Db 481 GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCACAGAGCCTGTGATCAGC 540
QY 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCAACAGTCAAAATATCTCTTGGG 600
Db 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCAACAGTCAAAATATCTCTTGGG 600
QY 601 ATGAAATGCCATATATATCCGGACCACTTGGATCTTTTATATTTATGATGATGTTT 660
Db 601 ATGAAATGCCATATATATCCGGACCACTTGGATCTTTTATATTTATGATGATGTTT 660
QY 661 GAAACATCAAGTCTGTGGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGGTC 720
Db 661 GAAACATCAAGTCTGTGGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGGTC 720
QY 721 CTTGTTCTTGTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGAGTTGTTCTTCTCT 780
```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1316	100.0	1797	3	US-09-875-811-11	Sequence 11, Appl
2	1316	100.0	1818	3	US-09-875-811-7	Sequence 7, Appl
3	1316	100.0	1866	3	US-09-875-811-3	Sequence 3, Appl
4	1316	100.0	1971	3	US-09-875-811-9	Sequence 9, Appl
5	1316	100.0	1992	3	US-09-875-811-5	Sequence 5, Appl
6	1316	100.0	2040	3	US-09-875-811-1	Sequence 1, Appl
7	1316	100.0	2420	3	US-09-875-811-13	Sequence 13, Appl
8	224.4	17.1	2858	3	US-09-949-016-5349	Sequence 5349, Ap
9	224.4	17.1	2882	2	US-08-424-567-1	Sequence 1, Appl
10	224.4	17.1	2882	2	US-08-711-928-1	Sequence 1, Appl
11	224.4	17.1	2882	3	US-09-184-937-1	Sequence 1, Appl
12	217.4	16.5	4927	3	US-09-949-016-5627	Sequence 5627, Ap
13	189.2	14.4	2441	3	US-09-785-381-4	Sequence 4, Appl
14	185	14.1	2832	3	US-09-949-016-4516	Sequence 4516, Ap
15	184	14.0	4113	3	US-09-785-381-2	Sequence 2, Appl
16	178.2	13.5	8774	3	US-09-949-016-16258	Sequence 16258, A
17	83	6.3	2913	3	US-09-795-927-6	Sequence 6, Appl
18	83	6.3	3749	3	US-09-795-927-8	Sequence 8, Appl
19	73.2	5.6	457	3	US-09-270-767-1395	Sequence 1395, Ap
20	73.2	5.6	457	3	US-09-270-767-16677	Sequence 16677, A
21	68.6	5.2	588	3	US-09-385-982-129	Sequence 129, App
22	67.8	5.2	623	3	US-09-385-982-513	Sequence 513, App
23	59.6	4.5	2022	3	US-09-720-317A-21	Sequence 21, Appl
24	52.6	4.0	1311	3	US-09-134-000C-545	Sequence 545, App





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:29:16 ; Search time 588 Seconds  
(without alignments)  
10564.825 Million cell updates/sec

Title: US-10-761-033-1-copy\_1\_1316

Perfect score: 1316  
Sequence: 1 atgcaggagcaagaggaa.....gtggataaaatcgattgggg 1316

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

EST:\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1291.2	98.1	5240	4 CR859003	CR859003 Pongo pyg
2	1124.8	85.5	2650	4 BC060784	BC060784 Homo sapi
3	1124.8	85.5	2650	4 BC060784	BC060784 Homo sapi
4	1069.6	81.3	4393	4 AK054327	AK054327 Mus muscu
5	962.4	73.1	1800	10 AY405555	AY405555 Homo sapi
6	959.2	72.9	1800	10 AY405556	AY405556 Pan trogl
7	762.4	57.9	1800	10 AY405557	AY405557 Mus muscu
8	631.4	48.0	645	3 BM726084	BM726084 UI-E-EJ0-
9	611.2	46.4	729	3 BM682518	BM682518 UI-E-EJ0-
10	562	42.7	792	3 CR622935	CR622935 DKFZp469J
11	556.2	42.3	799	6 CR745519	CR745519 UI-M-GV0-
12	524.4	39.8	773	6 CR764738	CR764738 DKFZp469M
13	518.6	39.4	744	7 CR762687	CR762687 DKFZp469P
14	515	39.1	864	6 CR992523	CR992523 AGENCOURT
15	497.4	37.8	725	7 CR549818	CR549818 DKFZp469J
16	469.8	35.7	715	6 CR744029	CR744029 UI-M-GV0-
17	467.8	35.5	787	7 CR769333	CR769333 DKFZp469P
18	447.8	34.0	689	7 CR791972	CR791972 DKFZp469G
19	443	33.7	656	7 CR767955	CR767955 DKFZp469M
20	442	33.6	656	7 CR771034	CR771034 DKFZp469A
21	438.2	33.3	768	6 CB989260	CB989260 AGENCOURT
22	426.2	32.4	939	2 BF143423	BF143423 601790065

23	424.2	32.2	638	5	BX484645	BX484645 DKFZp468N
24	407	30.9	670	2	BB666404	BB666404
25	373.2	28.4	875	8	CX342514	CX342514 JGI X2T46
26	373	28.3	827	7	CF993927	CF993927 AGENCOURT
27	362.8	27.6	928	6	CB993059	CB993059 AGENCOURT
28	358.8	27.3	568	7	CR774608	CR774608 DKFZp469I
29	352.8	26.8	562	7	CR771087	CR771087 DKFZp469K
30	347.2	26.4	593	7	CR557400	CR557400 DKFZp469I
31	345.8	26.3	555	7	CR751908	CR751908 DKFZp469O
32	336.6	25.6	533	7	CR554938	CR554938 DKFZp469N
33	336.2	25.5	548	7	CR753714	CR753714 DKFZp469J
34	331.8	25.2	578	7	CR753649	CR753649 DKFZp469H
35	319.2	24.3	658	7	CR790935	CR790935 DKFZp469J
36	300.8	22.9	524	6	CD803846	CD803846 UI-M-GV0-
37	299.4	22.8	418	6	CD701536	CD701536 EST18060
38	268	20.4	418	7	CR554171	CR554171 DKFZp469K
39	226.4	17.2	433	7	CR751437	CR751437 DKFZp469J
40	222.2	16.9	2140	11	DQ030554	DQ030554 Homo sapi
41	221	16.8	364	5	BY770037	BY770037
42	221	16.8	2624	4	HSM806949	BM640837 Homo sapi
43	220.6	16.8	2878	4	CR926479	CR926479 Pongo pyg
44	209.4	15.9	3603	4	AK052641	AK052641 Mus muscu
45	207.6	15.8	2140	11	DQ030555	DQ030555 Pan trogl

#### ALIGNMENTS

RESULT 1	CR859003	5240 bp	mrna	linear	HTC 17-APR-2005
LOCUS	Pongo pygmaeus mRNA; cDNA DKFZp469J151 (from clone DKFZp469J151).				
DEFINITION	CR859003				
ACCESSION	CR859003.1	GI:55728919			
VERSION	HTC.				
KEYWORDS	Pongo pygmaeus (orangutan)				
SOURCE	Pongo pygmaeus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pongo.				
REFERENCE	1 (bases 1 to 5240)				
AUTHORS	Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Well,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.				
CONSRM	The German cDNA Consortium				
TITLE	Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764				
JOURNAL	Neuherberg, GERMANY				
COMMENT	Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469J151) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp469J151 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
source	1..5240				
	/organism="Pongo pygmaeus"				
	/mol_type="mRNA"				
	/db_xref="taxon:9600"				
	/clone="DKFZp469J151"				
	/tissue="kidney"				
	/clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfiIa + SfiIb"				
	/dev_stage="adult"				
	/note="solute carrier family 26, member 7 isoform a (Homo sapiens)"				
gene	1..5240				
	/gene="DKFZp469J151"				
CDS	join(208..228,228..2177)				





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:22:12 ; Search time 724 Seconds  
(without alignments)

12114.276 Million cell updates/sec

Title: US-10-761-033-1\_COPY\_1\_1316

Perfect score: 1316

Sequence: 1 atgacaggagcaaaaggaa.....gtggataaaatcgattggggg 1316

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316	100.0	1797	6	ABK12981
2	1316	100.0	1818	6	ABK12979
3	1316	100.0	1866	6	ABK12977
4	1316	100.0	1971	6	ABK12980
5	1316	100.0	1971	10	ACF79733
6	1316	100.0	1992	6	ABK12978
7	1316	100.0	2040	6	ABK12976
8	1316	100.0	2345	10	ADC35055
9	1316	100.0	2420	6	ABK12982
10	1314.4	99.9	2631	5	ADL62785
11	1314.4	99.9	2886	10	ADC35054
12	1314.4	99.9	5265	6	ADC35215
13	1169.4	88.9	5312	13	ACN41251
14	1166.2	88.6	2848	6	ABK83213
15	1068	81.2	2289	10	ACF79734
16	514.4	39.1	1296	5	AS70486
17	224.4	17.1	2692	12	ADH22591
18	224.4	17.1	2879	9	ADA10954
19	224.4	17.1	2879	12	ADG47980

20	224.4	17.1	2882	2	AAQ71399	Aaq71399 Down-regu
21	224.4	17.1	2908	13	ACN43500	Acn43500 Human dia
22	224.4	17.1	2918	5	ADL62992	Adl62992 Human ova
23	217.4	16.5	2343	4	AAI99656	Aai99656 Human pen
24	217.4	16.5	2782	4	AAI99681	Aai99681 Mouse pen
25	217.4	16.5	4930	6	ABK92231	Abk92231 Prostate
26	217.4	16.5	4930	11	ADN39611	Adn39611 Cancer/an
27	217.4	16.5	4930	11	ADN95181	Adn95181 Human BSC
28	217.4	16.5	4930	12	ADJ74750	Adj74750 Human pen
29	217.4	16.5	4930	12	ADJ74895	Adj74895 Marker ge
30	214.2	16.3	2829	11	ADM56356	Adm56356 Frog SLC2
31	206.8	15.7	1835	8	ADA47351	Ada47351 Human tra
32	206.8	15.7	2378	13	ACN43826	Acn43826 Human dia
33	204	15.5	3049	10	ACF79737	Acf79737 Mouse ani
34	203.2	15.4	3635	11	ADM56348	Adm56348 Frog SLC2
35	201	15.3	249	5	ADL36859	Adl36859 Human ova
36	201	15.3	249	5	ADI71704	Adi71704 Human ova
37	201	15.3	274	5	ADL36868	Adl36868 Human ova
38	201	15.3	274	5	ADI71713	Adi71713 Human ova
39	201	15.3	373	5	ADL43243	Adl43243 Human ova
40	201	15.3	413	5	ADL43252	Adl43252 Human ova
41	199.4	15.2	2264	10	ADJ95139	Adj95139 Novel NOV
42	199.4	15.2	2264	11	ADM29327	Adm29327 Human nov
43	199	15.1	2869	13	ACN43501	Acn43501 Human dia
44	189.2	14.4	2441	11	ADM57788	Adm57788 Murine pr
45	189.2	14.4	2441	12	ADB86099	Ade86099 cDNA enco

## ALIGNMENTS

### RESULT 1

ABK12981

ID ABK12981 standard; cDNA; 1797 BP.

AC ABK12981;

DT 23-APR-2002 (first entry)

DB cDNA sequence of the gene encoding novel human protein (NHP) #6.

KW Novel human protein; NHP; transporter protein; polymorphism; gene;

KW mental disorder; biological disease; medical disorder; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..1797

FT /tag= a

FT /product= "Novel human protein (NHP) #6"

PN WO200194583-A2.

PD 13-DEC-2001.

PF 06-JUN-2001; 2001WO-US018393.

PR 07-JUN-2000; 2000US-0210045P.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Scoville J;

DR WPI; 2002-147673/19.

DR P-PSDB; AAU74917.

XX New human polynucleotides encoding proteins that share sequence

XX similarity with mammalian transporter proteins, useful for determining

XX genomic structures, identifying polymorphisms, or as reagents for

XX diagnosis or drug screening.

XX Disclosure; Page 45-46; 48pp; English.

PS

XX

CC The present invention relates to new isolated nucleic acid molecules  
CC comprising a nucleotide sequence encoding a protein that shares sequence  
CC similarity with mammalian transporter proteins. The invention also  
CC relates to a nucleotide sequence that hybridizes under stringent  
CC conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully  
CC defined in the specification or its complement. The isolated nucleic acid  
CC and the protein it encodes are useful for identifying a coding sequence  
CC and mapping a unique gene to a particular chromosome, identifying and  
CC characterizing the temporal and tissue specific expression of a gene,  
CC screening a human genomic library, determining the genomic structure of a  
CC given locus/allele or designing diagnostic tests. The nucleic acids and  
CC proteins are particularly useful for identifying polymorphisms and in  
CC amplification assays to detect mutations within the exons, introns and  
CC splice sites that can be used in diagnostics and pharmacogenomics. These  
CC are also useful for generating antibodies, as reagents in diagnostic  
CC assays or for identifying other cellular gene products related to novel  
CC human proteins. The nucleotides can be used as reagents in assays for  
CC screening for compounds that can be employed as pharmaceutical reagents  
CC useful in the therapeutic treatment of mental, biological or medical  
CC disorders and diseases. The present nucleic acid sequence encodes novel  
CC human protein #6 that is one of several (AU74912- AU74917) novel human  
CC proteins (NHP) of the invention  
XX

Sequence 1797 BP; 500 A; 367 C; 396 G; 534 T; 0 U; 0 Other;

Query Match 100.0%; Score 1316; DB 6; Length 1797;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACAGGAGCAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCAG 60  
Db 1 ATGACAGGAGCAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCAG 60  
Qy 61 TGTGAAGACATATACAGTGTGTAGAAAGGCGATCGCCCATTTTGGATGGGCAACAT 120  
Db 61 TGTGAAGACATATACAGTGTGTAGAAAGGCGATCGCCCATTTTGGATGGGCAACAT 120  
Qy 121 TACAATCTGAAGAAAGAACTTGCTCCAGACACTGTGTCTGGATAATGTGCGAGTTCAA 180  
Db 121 TACAATCTGAAGAAAGAACTTGCTCCAGACACTGTGTCTGGATAATGTGCGAGTTCAA 180  
Qy 181 CAGGTGACCCAGGATTTGGCCCTTTGCTGTTCTCTCATCTGTGCACCCAGTGTGGTTTA 240  
Db 181 CAGGTGACCCAGGATTTGGCCCTTTGCTGTTCTCTCATCTGTGCACCCAGTGTGGTTTA 240  
Qy 241 TATGGGTCTCTGTTTCTGCTCCATTAATTTATGCGATATTTGGAATGGGACATCATGTTGCC 300  
Db 241 TATGGGTCTCTGTTTCTGCTCCATTAATTTATGCGATATTTGGAATGGGACATCATGTTGCC 300  
Qy 301 ACAGGCACCTTTGCTTGACATCTTATATATCAGCCACGCGGTGGAACGGATTTGTCCT 360  
Db 301 ACAGGCACCTTTGCTTGACATCTTATATATCAGCCACGCGGTGGAACGGATTTGTCCT 360  
Qy 361 CAGAACATGCAAGATCTCACCACACAGAGTAACACGCGTGTGGGCTTATCCGACTTT 420  
Db 361 CAGAACATGCAAGATCTCACCACACAGAGTAACACGCGTGTGGGCTTATCCGACTTT 420  
Qy 421 GAAATGCAAGGATCCAGTGTGCGACAGTTCCTTCTTGGGAGGTGTGATTCAGGTG 480  
Db 421 GAAATGCAAGGATCCAGTGTGCGACAGTTCCTTCTTGGGAGGTGTGATTCAGGTG 480  
Qy 481 GCCATGTTTGTGTCGAATCGGCGAGTCCACATTTTGGTTCACAGAGCTGTGATCAGC 540  
Db 481 GCCATGTTTGTGTCGAATCGGCGAGTCCACATTTTGGTTCACAGAGCTGTGATCAGC 540  
Qy 541 GCAATGCAACATGCGGGCTGCCACCCATGTGTGTGACTTCACAAGTCAAAATATCTCTGGGA 600  
Db 541 GCAATGCAACATGCGGGCTGCCACCCATGTGTGTGACTTCACAAGTCAAAATATCTCTGGGA 600  
Qy 601 ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGATATGTTTTT 660  
Db 601 ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGATATGTTTTT 660

Qy 661 GAAAAACATCAAGTCTGTGCGAGCTGGAAGCATTTGCTTTTATCTTGTGAGCATTTGTGTC 720  
Db 661 GAAAAACATCAAGTCTGTGCGAGCTGGAAGCATTTGCTTTTATCTTGTGAGCATTTGTGTC 720  
Qy 721 CTTGTTCTTGTATAAGAGCTGAATGAACAGTTTAAAAAGGAAAAATTAAGTTGTTCTTCT 780  
Db 721 CTTGTTCTTGTATAAGAGCTGAATGAACAGTTTAAAAAGGAAAAATTAAGTTGTTCTTCT 780  
Qy 781 GTAGATTTAGTTTGTATTTGCTGCGATCATTTGCTTGTATTGCAACCAATATGGAAC 840  
Db 781 GTAGATTTAGTTTGTATTTGCTGCGATCATTTGCTTGTATTGCAACCAATATGGAAC 840  
Qy 841 ACATATGGAATAGAGTAGTTGGTCATATTCCACAGGAATTCCTCACCTAGAGCTCCC 900  
Db 841 ACATATGGAATAGAGTAGTTGGTCATATTCCACAGGAATTCCTCACCTAGAGCTCCC 900  
Qy 901 CCGATGAACATCTCTCTGCGGTGATCACTGAAGCTTTCCGAGTGGCACTTTGTAGGCTAT 960  
Db 901 CCGATGAACATCTCTCTGCGGTGATCACTGAAGCTTTCCGAGTGGCACTTTGTAGGCTAT 960  
Qy 961 GTGGCCTCAGTGGCTCTTGTCTCAAGGATCTGCCAAAAAATTCAAATATTCAATTGATGAC 1020  
Db 961 GTGGCCTCAGTGGCTCTTGTCTCAAGGATCTGCCAAAAAATTCAAATATTCAATTGATGAC 1020  
Qy 1021 AACACAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCAATTTTCTTCTGCAATA 1080  
Db 1021 AACACAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCAATTTTCTTCTGCAATA 1080  
Qy 1081 CCAAGTCTCTGTCATATCTTGCATTTTCTGCTTATAGTCTATGCAATAGGACCTTTG 1200  
Db 1081 CCAAGTCTCTGTCATATCTTGCATTTTCTGCTTATAGTCTATGCAATAGGACCTTTG 1200  
Qy 1201 CTTTACTGGCTGCCCATGTGTCTTTCAGAGCATTTATTTGTGGGACTGGAAGGAATG 1260  
Db 1201 CTTTACTGGCTGCCCATGTGTCTTTCAGAGCATTTATTTGTGGGACTGGAAGGAATG 1260  
Qy 1261 CTAATACAGTTCGAGATTTAAAAAATATTTGGAATGTGGAATAAATCGATTTGGG 1316  
Db 1261 CTAATACAGTTCGAGATTTAAAAAATATTTGGAATGTGGAATAAATCGATTTGGG 1316

## RESULT 2

ABK12979  
ID ABK12979 standard; cDNA; 1818 BP.

XX ABK12979;

XX 23-APR-2002 (first entry)

XX cDNA sequence of the gene encoding novel human protein (NHP) #4.

XX Novel human protein; NHP; transporter protein; polymorphism; gene;  
XX mental disorder; biological disease; medical disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1818

FT /\*tag= a

FT /product= "Novel human protein (NHP) #4"

XX WO200194583-A2.

XX 13-DEC-2001.

XX 06-JUN-2001; 2001WO-US018393.

XX 07-JUN-2000; 2000US-0210045P.

XX

No.	Score	Match	Length	DB	ID	Description
1	1316	100.0	1797	6	AR486412	Sequence
2	1316	100.0	1797	6	AX338603	Sequence
3	1316	100.0	1818	6	AR486410	Sequence
4	1316	100.0	1818	6	AX338599	Sequence
5	1316	100.0	1866	6	AR486408	Sequence
6	1316	100.0	1866	6	AX338595	Sequence
7	1316	100.0	1971	6	AR486411	Sequence
8	1316	100.0	1971	6	AX338601	Sequence
9	1316	100.0	1992	6	AR486409	Sequence
10	1316	100.0	1992	6	AX338597	Sequence
11	1316	100.0	2040	6	AR486407	Sequence
12	1316	100.0	2040	6	AX338593	Sequence
13	1316	100.0	2345	8	AX960423	Sequence
14	1316	100.0	2345	8	HS413230	Homo sapi
15	1316	100.0	2420	6	AR486413	Sequence
16	1316	100.0	2420	6	AX338605	Sequence
17	1316	100.0	2928	8	BC094730	Homo sapi
18	1316	100.0	5254	8	HS413228	Homo sapi

```
QY 241 TATGGGCTCTGTTTCTGSCCATATTTATGCCATATTTTGGAAATGGGACATCATGTTGCC 300
Db 241 TATGGGCTCTGTTTCTGSCCATATTTATGCCATATTTTGGAAATGGGACATCATGTTGCC 300
QY 301 ACAGGCACCTTTGCTTTCATCATCTTAAATATCAGCAACCGCGTGGAAATGTCCT 360
Db 301 ACAGGCACCTTTGCTTTCATCATCTTAAATATCAGCAACCGCGTGGAAATGTCCT 360
QY 361 CAGAACATGCAAGATCTCACCAACAGAGTAAACAGAGTGTGGCTTATCCGACTTT 420
Db 361 CAGAACATGCAAGATCTCACCAACAGAGTAAACAGAGTGTGGCTTATCCGACTTT 420
QY 421 GAAATGCAAGAGATCCAGTGTGTCAGCAGTTTCTTCTTGGAGAGTGTATTCAGGTG 480
Db 421 GAAATGCAAGAGATCCAGTGTGTCAGCAGTTTCTTCTTGGAGAGTGTATTCAGGTG 480
QY 481 GCCATGTTTGTGCTGCAACTGGGCAAGTCCCATTTTGGTCCACAGAGCCTGTATCAGC 540
Db 481 GCCATGTTTGTGCTGCAACTGGGCAAGTCCCATTTTGGTCCACAGAGCCTGTATCAGC 540
QY 541 GCAATGCAACTGGGGCTGCCACCATGTGTGACTTCAAGTCAATATCTCTTGGGA 600
Db 541 GCAATGCAACTGGGGCTGCCACCATGTGTGACTTCAAGTCAATATCTCTTGGGA 600
QY 601 ATGAAATGCCATATATATCCGACCACTTGGATTTCTTTTATTTATGTCATATGTTTT 660
Db 601 ATGAAATGCCATATATATCCGACCACTTGGATTTCTTTTATTTATGTCATATGTTTT 660
QY 661 GAAACATCAAGTCTGTGCGACTGGAAGCAITGCTTTTATCTTCTGAGCAATGTGTC 720
Db 661 GAAACATCAAGTCTGTGCGACTGGAAGCAITGCTTTTATCTTCTGAGCAATGTGTC 720
QY 721 CTGTTCTTGTAAAGAGCTGAATGAACAGTTTAAAGGAAAAATTAAGTTGTTCTTCT 780
Db 721 CTGTTCTTGTAAAGAGCTGAATGAACAGTTTAAAGGAAAAATTAAGTTGTTCTTCT 780
QY 781 GTAGATTTAGTTTATGTTATGTCATCATTTGCTTATTTGCAACCAATATGGAAC 840
Db 781 GTAGATTTAGTTTATGTTATGTCATCATTTGCTTATTTGCAACCAATATGGAAC 840
QY 841 ACATATGGAATAGAAATGTTGTCATATTCACCAAGGAAATTCCTCACCTAGAGCTCCC 900
Db 841 ACATATGGAATAGAAATGTTGTCATATTCACCAAGGAAATTCCTCACCTAGAGCTCCC 900
QY 901 CCGATGAACATCTCTCTGCGGTGATCATGAGCTTTGGAGTGGCACTTTGAGGCTAT 960
Db 901 CCGATGAACATCTCTCTGCGGTGATCATGAGCTTTGGAGTGGCACTTTGAGGCTAT 960
QY 961 GTGGCTCCTAGTGGCTTGTGCTCAAGGATCTGCCAAAAATTCAAATATTCATATGATGAC 1020
Db 961 GTGGCTCCTAGTGGCTTGTGCTCAAGGATCTGCCAAAAATTCAAATATTCATATGATGAC 1020
QY 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATGATTTTCTTCAATTTTCTTCTGCAATA 1080
Db 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATGATTTTCTTCAATTTTCTTCTGCAATA 1080
QY 1081 CCAAGTGTCTGTCATGGAAGGACGGCTGCTGTAACAGCAGGAGCGAAGACACAG 1140
Db 1081 CCAAGTGTCTGTCATGGAAGGACGGCTGCTGTAACAGCAGGAGCGAAGACACAG 1140
QY 1141 GTGGCTCTCTAATATCTTGCATTTTCTGCTTATAGTCAATATGCAATAGGACCTTTG 1200
Db 1141 GTGGCTCTCTAATATCTTGCATTTTCTGCTTATAGTCAATATGCAATAGGACCTTTG 1200
QY 1201 CTTTACTGGCTGCCCATGTGTCTTGCAGAGCAATTTTGTGGGAGTGAAGGGAATG 1260
Db 1201 CTTTACTGGCTGCCCATGTGTCTTGCAGAGCAATTTTGTGGGAGTGAAGGGAATG 1260
QY 1261 CTAATACAGTTCCGAGATTTAAAAAATAATTTGGAAATGGAATGGAATGGAATGGAATG 1316
Db 1261 CTAATACAGTTCCGAGATTTAAAAAATAATTTGGAAATGGAATGGAATGGAATGGAATG 1316
```

```
RESULT 2
AX338603 LOCUS AX338603 1797 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 11 from Patent WO0194583.
ACCESSION AX338603
VERSION AX338603.1 GI:18128993
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE 1
AUTHORS Walke,D.W. and Scoville,J.
TITLE Novel human transporter proteins and polynucleotides encoding the
JOURNAL Patent: WO 0194583-A 11 13-DEC-2001;
FEATURES Lexicon Genetics Incorporated (US)
            Location/Qualifiers
            source 1..1797
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1316; DB 6; Length 1797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACAGAGCAGCAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCCAG 60
Db 1 ATGACAGAGCAGCAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCCAG 60
QY 61 TGTGAAGACATTATACAGTGTGTAGAGGCGCATGCCCCATTTTGGATGGGACCAAT 120
Db 61 TGTGAAGACATTATACAGTGTGTAGAGGCGCATGCCCCATTTTGGATGGGACCAAT 120
QY 121 TACAACTCTGAAGAAACTTCTTCCAGACACCTGTCTGGGTAATGTTGGCGAGTTCAA 180
Db 121 TACAACTCTGAAGAAACTTCTTCCAGACACCTGTCTGGGTAATGTTGGCGAGTTCAA 180
QY 181 CAGGTGACCAAGGATTTGGCTTGTCTCTCATCTGTGACCCAGTGTGGTGT 240
Db 181 CAGGTGACCAAGGATTTGGCTTGTCTCTCATCTGTGACCCAGTGTGGTGT 240
QY 241 TATGGGCTCTGTTTCTGCGCATATTTATGCCATAATTTGGAATGGGACATCATGTTGCC 300
Db 241 TATGGGCTCTGTTTCTGCGCATATTTATGCCATAATTTGGAATGGGACATCATGTTGCC 300
QY 301 ACAGGCACCTTTGCTTTCATCATCTTAAATATCAGCAACCGCGTGGAAATGTCCT 360
Db 301 ACAGGCACCTTTGCTTTCATCATCTTAAATATCAGCAACCGCGTGGAAATGTCCT 360
QY 361 CAGAACATGCAAGATCTCACCAACAGAGTAAACAGAGTGTGGCTTATCCGACTTT 420
Db 361 CAGAACATGCAAGATCTCACCAACAGAGTAAACAGAGTGTGGCTTATCCGACTTT 420
QY 421 GAAATGCAAGAGATCCAGTGTGTCAGCAGTTTCTTCTTGGAGAGTGTATTCAGGTG 480
Db 421 GAAATGCAAGAGATCCAGTGTGTCAGCAGTTTCTTCTTGGAGAGTGTATTCAGGTG 480
QY 481 GCCATGTTTGTGCTGCAACTGGGCAAGTCCCATTTTGGTCCACAGAGCCTGTATCAGC 540
Db 481 GCCATGTTTGTGCTGCAACTGGGCAAGTCCCATTTTGGTCCACAGAGCCTGTATCAGC 540
QY 541 GCAATGCAACTGGGGCTGCCACCATGTGTGACTTCAAGTCAATATCTCTTGGGA 600
Db 541 GCAATGCAACTGGGGCTGCCACCATGTGTGACTTCAAGTCAATATCTCTTGGGA 600
QY 601 ATGAAATGCCATATATATCCGACCACTTGGATTTCTTTTATTTATGTCATATGTTTT 660
Db 601 ATGAAATGCCATATATATCCGACCACTTGGATTTCTTTTATTTATGTCATATGTTTT 660
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:49:38 ; Search time 24 Seconds  
(without alignments)  
509.146 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2332  
Sequence: 1 MTGAKRKKKMLWSKMTQP.....MLIQFRDLKKYNNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA New:
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	805	36.1	735	6	US-10-505-263-89
2	774	34.7	739	6	US-10-505-263-12
3	753.5	33.8	739	6	US-10-505-263-96
4	742.5	33.3	719	6	US-10-505-263-87
5	734	32.9	753	6	US-10-505-263-91
6	732	32.8	738	6	US-10-505-263-4
7	732	32.8	758	6	US-10-505-263-2
8	725.5	32.5	735	6	US-10-505-263-8
9	725.5	32.5	758	6	US-10-505-263-6
10	691.5	31.0	778	6	US-10-505-263-81
11	686.5	30.8	769	6	US-10-505-263-83
12	685	30.7	788	6	US-10-505-263-85
13	654	29.3	704	6	US-10-505-263-10
14	645.5	28.9	791	6	US-10-537-002-65
15	431	19.3	616	7	US-11-096-568A-33957
16	431	19.3	685	7	US-11-096-568A-33956
17	431	19.3	717	7	US-11-096-568A-33955
18	418	18.7	93	6	US-10-505-263-76
19	396	17.7	634	7	US-11-096-568A-31697
20	396	17.7	669	7	US-11-096-568A-31696
21	392	17.6	658	7	US-11-096-568A-703
22	369.5	16.6	677	7	US-11-010-239-48
23	366.5	16.4	560	7	US-11-096-568A-31698
24	357	16.0	649	7	US-11-010-239-44
25	285.5	12.8	518	7	US-11-096-568A-704

ALIGNMENTS

RESULT 1  
US-10-505-263-89  
; Sequence 89, Application US/10505263  
; Publication No. US20060014940A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Case Western Reserve University  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Mount, David B  
; APPLICANT: Romero, Michael  
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2  
; FILE REFERENCE: 1242/50/2 PCT/US  
; CURRENT APPLICATION NUMBER: US/10/505,263  
; CURRENT FILING DATE: 2004-08-20  
; PRIOR APPLICATION NUMBER: US 60/360,275  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: PCT/US03/06469  
; PRIOR FILING DATE: 2003-02-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 89  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Xenopus laevis  
US-10-505-263-89

Query Match 36.1%; Score 805; DB 6; Length 735;  
Best Local Similarity 39.2%; Pred. No. 6.1e-56;  
Matches 159; Conservative 91; Mismatches 154; Indels 2; Gaps 1;

Qy	32	LPIDWAPHYKLLPDTVSGIMLAVQVQVQGLAPAVLSSVHPVFGYGLSLPPAIYA	91
Db	67	IPICGLPFPYVKWLLGDIVSLGVLIIQPLQGLAVALLAGVPVFGYGLSFPFVLIIYA	126
Qy	92	IFGMGHVATGTFTALTSLSANAVRIVPQNMQLTTQSNTSVLGLSDPFEMQRIHVA	151
Db	127	IFGTSRHISPTFAVISVMVGSVTESLVPS--ENVRLPGNESVIDIARDNDREVASAL	184
Qy	152	SFLGGVQVAMFVLQGSATFVVTPEVISAMTGAATHVTSQVKYLLGMKMPYISG	211
Db	185	TFLVGLFQIMLGLVQGVFVVTYLSPLTIRGTYSAAAIHVTVSQMKSVLGVQIS	244
Qy	212	PFYIYAVFENIKSVRLKALLSLSLVVLVKELNQSFKRKIVLVPDLVLIIAASF	271
Db	245	LIYAFVNLCAKPTNTASLLGICISITVLPFLNPKNDKYSKIMPIELITIVATG	304
Qy	272	ACYCTNMNTYGLEVVGVHPIQGIPIPSRAPPMNLSAVITEAFGVALVGVYVSLA	331

Db	305	ISYGASLHQVGV	IVGSLPTGKAPMLPNTNIFARVVGNAFAIAVWVYAF	TISLAKMEG	364	
Qy	332	KKFKYSIDNDOEFLAHGLSNIVSSPFFCIPSAAMGR	TAGLSYTGAKTQVACLLISCFVL	391		
Db	365	VKHGYNIDNQELIALGLSNISGSPFQCTTG	TMSRSLVQBSTGGHSAVASLSVL	424		
Qy	392	IVIVAIGPLLWLMCMVLASII	VGLKGMLIQFRDLKKYVNV	DKID	437	
Db	425	IIILKAGELFETLPKAI	LAAYVVNLKGIYQFTDVPML	WRSNKFD	470	
RESULT 2						
US-10-505-263-12						
; Sequence 12, Application US/10505263						
; Publication No. US20060014940A1						
; GENERAL INFORMATION:						
; APPLICANT: Vanderbilt University						
; APPLICANT: Case Western Reserve University						
; APPLICANT: The Brigham and Women's Hospital, Inc.						
; APPLICANT: Mount, David B						
; APPLICANT: Romero, Michael						
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and S1						
; TITLE OF INVENTION: ANION EXCHANGERS						
; FILE REFERENCE: 1242/50/2 PCT/US						
; CURRENT APPLICATION NUMBER: US/10/505,263						
; CURRENT FILING DATE: 2004-08-20						
; PRIOR APPLICATION NUMBER: US 60/360,275						
; PRIOR FILING DATE: 2002-02-28						
; PRIOR APPLICATION NUMBER: PCT/US03/06469						
; PRIOR FILING DATE: 2003-02-28						
; NUMBER OF SEQ ID NOS: 96						
; SOFTWARE: PatentIn version 3.2						
; SEQ ID NO 12						
; LENGTH: 739						
; TYPE: PRT						
; ORGANISM: Mus musculus						
US-10-505-263-12						
Query Match 34.7%; Score 774; DB 6; Length 739;						
Best Local Similarity 33.1%; Pred. No. 1.7e-53;						
Matches 155; Conservative 108; Mismatches 173; Indels 32; Gaps 2						
Qy	2	TKAKRKKKMLSKWNGHTPQCE	DIQWCRRL-----PILDWAPHYN	42		
Db	44	TGHRRAFRRTHMLERKPD	TDIKQFVIRELQKSCSAKVRDCAFDPFVLRWLPKYD	103		
Qy	43	LKENLLDPTVSGIMLAVQOVTQGLAF	AVLSVHPVFGLYGSLFPATII	VAIFQMGHHVATG	102	
Db	104	LKKNILGDVMSGLVIGILLVPQSI	AYSLLAGOEPIYGLTSPFASII	YFLFCTSRHSIVG	163	
Qy	103	TFALTSLISANAVRIPQNNQLTQS	-----NTSVLGHSDPEMQRIHVA	149		
Db	164	IFGLCLMGIVGORELHKACPD	TDATSSIAVPSSGCVVNHNTDGLCDKSKCYAIKGS	223		
Qy	150	AVSPFLGGYIQVAMFVLQGSATF	VVTEPVIAMITGAATHVVTQVKLLGKMPFI	ISGP	209	
Db	224	TVTWMAGYQVAMGPFQVGVSV	LYSDALLSGFTVGTASFTILTQAKYLLGLSLPRSHGV	283		
Qy	210	LGFPYIYAVPENTKSVRL	EALLLSLISIVLVVLKELNEQPKRKIKVVL	PVDVLVLI	269	
Db	284	GSVITTTWTHIPRINTRN	TNICDLTSLCLLVLPKSELNEHPKDKAKIP	VELIIVVVA	343	
Qy	270	SPACYCTNMNTYGLEVVGHI	POGTPSPRAPMNTLSAVITEAFGVALVGVYASL	LAQ	329	
Db	344	TLASHFGKLGNGYNSIAGH	ITPTGMPKPAQDWSLIPNVA	VDIAISIIIGFAITVSLSEM	403	
Qy	330	SAKLPKYSIDNQBF	LAHGLSNIVSSPFFCIPSAAMGR	TAGLSYTGAKTQVACLLISCF	389	
Db	404	PAKHGTVKANQEM	YAGFCNII	PSPFHCITTSAAKTLVKES	TGCGTQLSAIVTALV	463
Qy	390	VLIVIVAIGPLLWLMCMVL	ASII	VGLKGMLIQFRDLKKYVNV	DKID	437



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:48:48 ; Search time 168 Seconds  
(without alignments)  
1091.828 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2232  
Sequence: 1 MTGAKRKKKSLMSKMTTPQ.....MLIQRLDKKYNVDKIDWG 439

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2232	100.0	598	3	US-09-875-811-12
2	2232	100.0	598	4	US-10-761-033-12
3	2232	100.0	605	3	US-09-875-811-8
4	2232	100.0	605	4	US-10-761-033-8
5	2232	100.0	621	3	US-09-875-811-4
6	2232	100.0	621	4	US-10-761-033-4
7	2232	100.0	656	3	US-09-875-811-10
8	2232	100.0	656	4	US-10-761-033-10
9	2232	100.0	663	3	US-09-875-811-6
10	2232	100.0	663	4	US-10-761-033-6
11	2232	100.0	663	4	US-10-222-009-14
12	2232	100.0	663	4	US-10-761-033-6
13	2232	100.0	679	3	US-09-875-811-2
14	2232	100.0	679	4	US-10-761-033-2
15	2227	99.8	656	4	US-10-222-009-13
16	2077	93.1	656	5	US-10-505-316-2
17	2077	93.1	656	5	US-10-505-316-4
18	2064.5	92.5	671	4	US-10-415-378-4
19	938.5	42.0	431	5	US-10-450-763-36658
20	825.5	37.0	555	4	US-10-478-758-6
21	753.5	33.8	739	4	US-10-205-823-369
22	753.5	33.8	739	4	US-10-328-194A-3
23	753.5	33.8	739	4	US-10-736-461-1
24	753.5	33.8	739	6	US-11-051-454-369
25	747.5	33.5	780	5	US-10-631-467-3
26	747.5	33.5	780	5	US-10-631-467-667
27	747.5	33.5	790	4	US-10-295-027-930

28	741.5	33.2	764	3	US-09-981-353-73	Sequence 73, Appl
29	741.5	33.2	764	4	US-10-235-994-20	Sequence 20, Appl
30	732	32.8	714	3	US-09-749-589-4	Sequence 4, Appl
31	732	32.8	714	5	US-10-684-532-4	Sequence 4, Appl
32	721.5	32.3	751	3	US-09-795-633-14	Sequence 14, Appl
33	721.5	32.3	751	4	US-10-156-239-14	Sequence 14, Appl
34	721.5	32.3	751	4	US-10-199-485-14	Sequence 14, Appl
35	721.5	32.3	780	4	US-10-420-495-11	Sequence 11, Appl
36	721.5	32.3	4115	4	US-10-038-854-4	Sequence 4, Appl
37	718.5	32.2	780	5	US-10-631-467-14	Sequence 14, Appl
38	718.5	32.2	780	5	US-10-631-467-1456	Sequence 1456, Ap
39	708	31.7	744	4	US-10-420-495-1	Sequence 1, Appl
40	708	31.7	744	4	US-10-420-495-3	Sequence 3, Appl
41	702	31.5	505	4	US-10-467-685-4	Sequence 4, Appl
42	662.5	29.7	790	5	US-10-505-316-10	Sequence 10, Appl
43	645.5	28.9	753	5	US-09-749-589-2	Sequence 2, Appl
44	645.5	28.9	753	5	US-10-684-532-2	Sequence 2, Appl
45	645.5	28.9	791	5	US-10-505-316-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-09-875-811-12  
; Sequence 12, Application US/09875811  
; Patent No. US20020032321A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. US20020032321A1 Human Transporter Proteins and Polynucleotid  
; FILE OF INVENTION: Same  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; CURRENT FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 598  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-875-811-12

Query Match 100.0%; Score 2232; DB 3; Length 598;  
Best Local Similarity 100.0%; Pred. No. 8.5e-201; Mismatches 0; Indels 0; Gaps 0;  
Matches 439; Conservative 0

Qy	1	MTGAKRKKKSLMSKMTTPQCE	IIQMCRRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ	60
Db	1	MTGAKRKKKSLMSKMTTPQCE	IIQMCRRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ	60
Qy	61	QVTGGLAPAVLSVHPVPGLYGSLFP	PAIIYAFGNGHVAITGTFTALTSLISANAVERIVP	120
Db	61	QVTGGLAPAVLSVHPVPGLYGSLFP	PAIIYAFGNGHVAITGTFTALTSLISANAVERIVP	120
Qy	121	QNMQLTQSTNSVLGLSDPEMRQIHV	AAAVFLGGVIOQAMFVLQGSATFVVTPEVIS	180
Db	121	QNMQLTQSTNSVLGLSDPEMRQIHV	AAAVFLGGVIOQAMFVLQGSATFVVTPEVIS	180
Qy	181	AMTTGAATHVVTQVKLLGKMPYIS	GPLGFFFIYAYVFNKSVRLLEALLLSLSIW	240
Db	181	AMTTGAATHVVTQVKLLGKMPYIS	GPLGFFFIYAYVFNKSVRLLEALLLSLSIW	240
Qy	241	LVLVKELNEQPKRKIKVVLVDVL	LVIIAASFACVTNNNTYGLGVGHIPOGIPSPRAP	300
Db	241	LVLVKELNEQPKRKIKVVLVDVL	LVIIAASFACVTNNNTYGLGVGHIPOGIPSPRAP	300
Qy	301	PMNITLSAVITEAFGVALGVYAS	LALAQGSAKPKYSIDDDNQEFLAHGLSNIVSSPFPCI	360
Db	301	PMNITLSAVITEAFGVALGVYAS	LALAQGSAKPKYSIDDDNQEFLAHGLSNIVSSPFPCI	360

Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Qy 421 LIQPRDLKKYNNVDKIDWG 439  
Db 421 LIQPRDLKKYNNVDKIDWG 439

## RESULT 2

US-10-761-033-12  
; Sequence 12, Application US/10761033  
; Publication No. US20040138442A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/10/761,033  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US/09/875,811  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 598  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-761-033-12

Query Match 100.0%; Score 2232; DB 4; Length 598;  
Best Local Similarity 100.0%; Pred. No. 8.5e-201;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60  
Db 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60  
Qy 61 QVTOGLAPAVLSSVHPVFLYGLSLFPALIIYAFMGHHVATGTFTALTSLISANAVRIVP 120  
Db 61 QVTOGLAPAVLSSVHPVFLYGLSLFPALIIYAFMGHHVATGTFTALTSLISANAVRIVP 120  
Qy 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Db 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLLEALLLSLSIVV 240  
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLLEALLLSLSIVV 240  
Qy 241 LVLVKELNEQPKRKIKVLPVDLVIIAASPCYCTNMNTYGLGVGHIPOGIPSPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVIIAASPCYCTNMNTYGLGVGHIPOGIPSPRAP 300  
Qy 301 PMNLSAVITEAFGVALGVYASLALAQSAKPKYSIDDDNOEFLAHGLSNIVSFPFCI 360  
Db 301 PMNLSAVITEAFGVALGVYASLALAQSAKPKYSIDDDNOEFLAHGLSNIVSFPFCI 360  
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Qy 421 LIQPRDLKKYNNVDKIDWG 439  
Db 421 LIQPRDLKKYNNVDKIDWG 439

## RESULT 3

US-09-875-811-8  
; Sequence 8, Application US/09875811

; Patent No. US20020032321A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. US20020032321A1el Human Transporter Proteins and Polynucleotides  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; CURRENT FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-875-811-8

Query Match 100.0%; Score 2232; DB 3; Length 605;  
Best Local Similarity 100.0%; Pred. No. 8.7e-201;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60  
Db 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60  
Qy 61 QVTOGLAPAVLSSVHPVFLYGLSLFPALIIYAFMGHHVATGTFTALTSLISANAVRIVP 120  
Db 61 QVTOGLAPAVLSSVHPVFLYGLSLFPALIIYAFMGHHVATGTFTALTSLISANAVRIVP 120  
Qy 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Db 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLLEALLLSLSIVV 240  
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLLEALLLSLSIVV 240  
Qy 241 LVLVKELNEQPKRKIKVLPVDLVIIAASPCYCTNMNTYGLGVGHIPOGIPSPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVIIAASPCYCTNMNTYGLGVGHIPOGIPSPRAP 300  
Qy 301 PMNLSAVITEAFGVALGVYASLALAQSAKPKYSIDDDNOEFLAHGLSNIVSFPFCI 360  
Db 301 PMNLSAVITEAFGVALGVYASLALAQSAKPKYSIDDDNOEFLAHGLSNIVSFPFCI 360  
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIAIGPLLYWLPVCVLASIIIVVGLKGM 420  
Qy 421 LIQPRDLKKYNNVDKIDWG 439  
Db 421 LIQPRDLKKYNNVDKIDWG 439

## RESULT 4

US-10-761-033-8  
; Sequence 8, Application US/10761033  
; Publication No. US20040138442A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/10/761,033  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US/09/875,811  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:48:03 ; Search time 46 Seconds  
(without alignments)  
789.014 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2232  
Sequence: 1 MTGAKRKKKSMKMTPOC.....MLIQFRLKXYWVNDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/PCUS\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/BE\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2232	100.0	598	2	US-09-875-811-12
2	2232	100.0	605	2	US-09-875-811-8
3	2232	100.0	621	2	US-09-875-811-4
4	2232	100.0	656	2	US-09-875-811-10
5	2232	100.0	663	2	US-09-875-811-6
6	2232	100.0	679	2	US-09-875-811-2
7	753.5	33.8	748	2	US-09-949-016-10387
8	747.5	33.5	803	2	US-09-949-016-11498
9	741.5	33.2	764	1	US-08-424-567-2
10	741.5	33.2	764	1	US-08-711-928-2
11	741.5	33.2	764	2	US-09-184-937-2
12	741.5	33.2	790	2	US-09-949-016-11220
13	721.5	32.3	780	2	US-09-785-381-11
14	708	31.7	744	2	US-09-785-381-1
15	708	31.7	744	2	US-09-785-381-3
16	472	21.1	593	2	US-09-720-317A-22
17	455.5	20.4	621	2	US-09-720-317A-16
18	453	20.3	970	2	US-09-795-927-7
19	451.5	20.2	646	2	US-09-720-317A-28
20	444.5	19.9	680	2	US-09-720-317A-18
21	443.5	19.9	658	2	US-09-720-317A-24
22	423	19.0	685	2	US-09-720-317A-31
23	421	18.9	295	2	US-09-785-381-5
24	418.5	18.8	646	2	US-09-720-317A-26
25	414	18.5	579	2	US-09-720-317A-4
26	412.5	18.5	660	2	US-09-720-317A-23
27	408	18.3	688	2	US-09-720-317A-2

28	395	17.7	656	2	US-09-720-317A-20	Sequence 20, Appl
29	389.5	17.5	590	2	US-09-902-540-14944	Sequence 14944, A
30	389.5	17.5	631	2	US-09-720-317A-29	Sequence 29, Appl
31	388	17.4	644	2	US-09-720-317A-25	Sequence 25, Appl
32	383	17.2	596	2	US-09-252-991A-23812	Sequence 23812, A
33	376	16.8	660	2	US-09-720-317A-30	Sequence 30, Appl
34	351	15.7	585	2	US-09-602-787A-616	Sequence 616, App
35	322	14.4	828	2	US-09-248-796A-20746	Sequence 20746, A
36	306	13.7	616	2	US-09-543-681A-4421	Sequence 4421, Ap
37	297	13.3	566	2	US-09-543-681A-4544	Sequence 4544, Ap
38	282.5	12.7	533	2	US-09-107-532A-5244	Sequence 5244, Ap
39	282.5	12.7	537	2	US-09-602-787A-618	Sequence 618, App
40	271	12.1	562	2	US-09-489-039A-10405	Sequence 10405, A
41	263	11.8	436	2	US-09-134-000C-3950	Sequence 3950, Ap
42	263	11.8	535	2	US-09-252-991A-21805	Sequence 21805, A
43	255.5	11.4	483	2	US-09-710-279-3132	Sequence 3132, Ap
44	250	11.2	575	2	US-09-438-185A-1015	Sequence 1015, Ap
45	243	10.9	151	2	US-09-270-767-33024	Sequence 33024, A

ALIGNMENTS

RESULT 1

US-09-875-811-12  
; Sequence 12, Application US/09875811  
; Patent No. 6703495  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; CURRENT FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 598  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-875-811-12

Query Match	100.0%;	Score	2232;	DB 2;	Length	598;			
Best Local Similarity	100.0%;	Pred. No.	2.1e-215;	Mismatches	0;	Indels	0;	Gaps	0;
Matches	439;	Conservative	0;						
Qy	1	MTGAKRKKKSMKSWKMTPOCEDIIQWCRRLPLTDWAPHYLNKENLLPDTVSGIMLAVQ	60						
Db	1	MTGAKRKKKSMKSWKMTPOCEDIIQWCRRLPLTDWAPHYLNKENLLPDTVSGIMLAVQ	60						
Qy	61	QVTOGLAPAVLSSVHPVFGYLSFLPAIIYAIKGMGHVATGTPALTSLISANAVRIVP	120						
Db	61	QVTOGLAPAVLSSVHPVFGYLSFLPAIIYAIKGMGHVATGTPALTSLISANAVRIVP	120						
Qy	121	QNMQLTQNTSVLGLSDPEMQRIHVAAVSVFLGGVIQVAMFVLQGSATFVVTPEVIS	180						
Db	121	QNMQLTQNTSVLGLSDPEMQRIHVAAVSVFLGGVIQVAMFVLQGSATFVVTPEVIS	180						
Qy	181	AMTTGAATHVVTQVKYLLGKMPYISGPIGFFYIYAVFENIKSVRLRALLLSLSTV	240						
Db	181	AMTTGAATHVVTQVKYLLGKMPYISGPIGFFYIYAVFENIKSVRLRALLLSLSTV	240						
Qy	241	LVLVKELNEQPKRIKVVLPVDLIIIAASFACVCTNMNTYGLVGVCHIQGIPSPAP	300						
Db	241	LVLVKELNEQPKRIKVVLPVDLIIIAASFACVCTNMNTYGLVGVCHIQGIPSPAP	300						
Qy	301	PMNLTSAVITEAFGVAVGVYASLALAQGSACKPKYSIDDDNQEFHLAHLGSLNIVSFFPCI	360						
Db	301	PMNLTSAVITEAFGVAVGVYASLALAQGSACKPKYSIDDDNQEFHLAHLGSLNIVSFFPCI	360						

QY 361 PSAAAMGRTAGLSTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLMPCVCLASIIIVVGLKGM 420  
Db ||||||||||||||||||  
QY 421 LIQFRDLKKYNNVDKIDWG 439  
Db ||||||||||||||||||  
QY 421 LIQFRDLKKYNNVDKIDWG 439

## RESULT 2

US-09-875-811-8  
; Sequence 8, Application US/09875811  
; Patent No. 6703495  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-875-811-8

Query Match 100.0%; Score 2232; DB 2; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.1e-215;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60  
Db 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60  
QY 61 QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGHHVATGTALTSLISANAVERIVP 120  
Db 61 QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGHHVATGTALTSLISANAVERIVP 120  
QY 121 QNMQLTQNTSVLGLSDFEMQRIHVAAVSPFLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Db 121 QNMQLTQNTSVLGLSDFEMQRIHVAAVSPFLGGVIVQAMFVLQGSATFVVTPEVIS 180  
QY 181 AMTTGAATHVTSQVYLLGKMPYISGPLGFFIYAYVFNENIKSVRLKALLLSISIV 240  
Db 181 AMTTGAATHVTSQVYLLGKMPYISGPLGFFIYAYVFNENIKSVRLKALLLSISIV 240  
QY 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300  
QY 301 PMNILSAVITEAFGVVAVSLAQAQSAKKPKYSIDNQBFLAHLGSLNIVSSPPFCI 360  
Db 301 PMNILSAVITEAFGVVAVSLAQAQSAKKPKYSIDNQBFLAHLGSLNIVSSPPFCI 360  
QY 361 PSAAAMGRTAGLSTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLMPCVCLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLSTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLMPCVCLASIIIVVGLKGM 420  
QY 421 LIQFRDLKKYNNVDKIDWG 439  
Db 421 LIQFRDLKKYNNVDKIDWG 439

## RESULT 3

US-09-875-811-4  
; Sequence 4, Application US/09875811  
; Patent No. 6703495  
; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 621  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-875-811-4

Query Match 100.0%; Score 2232; DB 2; Length 621;  
Best Local Similarity 100.0%; Pred. No. 2.2e-215;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60  
Db 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60  
QY 61 QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGHHVATGTALTSLISANAVERIVP 120  
Db 61 QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGHHVATGTALTSLISANAVERIVP 120  
QY 121 QNMQLTQNTSVLGLSDFEMQRIHVAAVSPFLGGVIVQAMFVLQGSATFVVTPEVIS 180  
Db 121 QNMQLTQNTSVLGLSDFEMQRIHVAAVSPFLGGVIVQAMFVLQGSATFVVTPEVIS 180  
QY 181 AMTTGAATHVTSQVYLLGKMPYISGPLGFFIYAYVFNENIKSVRLKALLLSISIV 240  
Db 181 AMTTGAATHVTSQVYLLGKMPYISGPLGFFIYAYVFNENIKSVRLKALLLSISIV 240  
QY 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300  
QY 301 PMNILSAVITEAFGVVAVSLAQAQSAKKPKYSIDNQBFLAHLGSLNIVSSPPFCI 360  
Db 301 PMNILSAVITEAFGVVAVSLAQAQSAKKPKYSIDNQBFLAHLGSLNIVSSPPFCI 360  
QY 361 PSAAAMGRTAGLSTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLMPCVCLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLSTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLMPCVCLASIIIVVGLKGM 420  
QY 421 LIQFRDLKKYNNVDKIDWG 439  
Db 421 LIQFRDLKKYNNVDKIDWG 439

## RESULT 4

US-09-875-811-10  
; Sequence 10, Application US/09875811  
; Patent No. 6703495  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/09/875,811  
; CURRENT FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 656  
; TYPE: PRT

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:44:08 ; Search time 40 seconds  
(without alignments)  
1055.979 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2232  
Sequence: 1 MTGAKRKKXKMLMSRHTPQ.....MLIQPRDLKRYWVDKIDWG 439

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	752.5	33.7	739	2 A54808	diastrophic dyspla
2	741.5	33.2	764	2 A47456	down-regulated in
3	668.5	30.0	703	2 A49994	sulfate transporte
4	648	29.0	611	2 T27820	hypothetical prote
5	638.5	28.6	700	2 T23629	hypothetical prote
6	611	27.4	749	2 T23628	hypothetical prote
7	571.5	25.6	758	2 T25751	hypothetical prote
8	553.5	24.8	380	2 T26165	hypothetical prote
9	543	24.3	650	2 T16077	hypothetical prote
10	496.5	22.2	750	2 B90137	sulfate permease
11	486.5	21.8	809	2 T32945	hypothetical prote
12	451.5	20.2	646	2 T01079	sulfate transport
13	447.5	20.0	605	2 G70409	high affinity sulf
14	443.5	19.9	658	2 T49069	sulfate transporte
15	443.5	19.9	658	2 T48901	sulfate transporte
16	443.5	19.9	703	2 T51161	hypothetical prote
17	418.5	18.8	646	2 T48902	sulfate transporte
18	412.5	18.5	553	2 A11498	probable sulfate t
19	409.5	18.3	553	2 AB1140	probable sulfate t
20	409.5	18.3	662	2 S51764	sulfate transport
21	403.5	18.1	667	2 S51763	sulfate transport
22	393.5	17.6	631	2 B86365	probable sulphate
23	388	17.4	644	2 S51765	sulfate transport
24	386	17.3	660	2 T04416	sulfate transport
25	372	16.7	566	2 S74633	high affinity sulf
26	371.5	16.6	573	2 D83440	probable sulfate t
27	369.5	16.6	677	2 T50022	sulfate transporte
28	357	16.0	649	2 T51839	sulfate transport
29	357	16.0	649	2 T00946	probable sulfate t

ALIGNMENTS

RESULT 1

A54808  
diastrophic dysplasia-associated sulfate transport protein - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 07-May-1999  
C:Accession: A54808  
R:Haestbacka, J.; de la Chapelle, A.; Mahtani, M.M.; Clines, G.; Reeve-Daly, M.P.; Daly, der, E.S.  
Cell, 78, 1073-1087, 1994  
A:Title: The diastrophic dysplasia gene encodes a novel sulfate transporter: positional c  
A:Reference number: A54808; MUID:95007757; PMID:7923357  
A:Accession: A54808  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-739 <HAB>  
A:Cross-references: UNIPARC:UPI0000178906; GB:U14528  
C:Genetics:  
A:Gene: GDB:DTD; DTDST  
A:Cross-references: GDB:125421; OMIM:222600  
A:Map position: 5q32-5q33.1  
C:Superfamily: sulfate transport protein

Query Match 33.7%; Score 752.5; DB 2; Length 739;

Best Local Similarity 35.6%; Pred. No. 8.9e-51;  
Matches 149; Conservative 98; Mismatches 159; Indels 13; Gaps 2;

Qy	32	LPILDWAPHNLKENLLPDTVSGIMLAVQVQTQGLAPVLSVHPVGLYGLSPALIYA	91
Db	93	LPVLQMLPKYDLKKNILGDMVSLIVGILLVPOSIAYSLLAGQBPVGLYTSFPFASIYF	152
Qy	92	IFGMGHVATGTPTALTSLSANAVRIVPO-----NMQNLTTQSNLSVL-----GLS	138
Db	153	LLGTSRHSIVGIFVLCMLGETVDRELQAGYDHAHSAPSLGMVNSGSLNHTSDRIC	212
Qy	139	DFEMQRIHVAASVFLGGVIQVAMFVLQGSATPVVTEPVVISAMTTGAATHVTSQVKYL	198
Db	213	DKSCYALMVGSTVTFIAGVTVQVAMGFQGVSVYLSALLSGFPVTGTLTSQAKYL	272
Qy	199	LGKMPYISGLPFPFYIYAVFENIKSVRLKLLLSIVLVVKELNEQPKRKIKV	258
Db	273	LGLNLPRTNGVSLTITTHVFRNIHKTNIICDLTSLCLLVLLPTKELNHFHKSCLKAP	332
Qy	259	LPVDVLIIAASPACYCTNMNTYGLRVGHHIPOGISPRAPPNNILSAVITFAFGVALV	318
Db	333	IPILVVVVAATLASHFGKLUHNTYNSIAGHIPGFMPPKVPENWLIPSAVDAIAISII	392
Qy	319	GYVASLALAQSAKKPKYISIDDDNQEFLLAGHLSNITVSSFFFCIPSAAMGRTAGLYSTGAK	378
Db	393	GFAITVLSLSEMFAXKHGVTVKANQEVVAIGPCNIIPSFHCFHTTSAALAKTLVKESTGCH	452
Qy	379	TQVACLISCFLVLIIVVIAIGPLLIWLMPCVLASIIIVGLKGLIQPRDLKRYWVDKID	437





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OH protein - protein search, using sw model

Run on: March 9, 2006, 18:40:53 ; Search time 232 Seconds  
(without alignments)  
1335.030 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2332

Sequence: 1 MTGAKRKKKSLMKSMTPO.....MLIQRLDKYWNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2232	100.0	656	2 Q96RN2_HUMAN	Q96rn2 homo sapien
2	2232	100.0	663	2 Q8TE53_HUMAN	Q8te53 homo sapien
3	2227	99.8	656	2 Q8TE54_HUMAN	Q8te54 homo sapien
4	2208	98.9	656	2 Q8RAL2_PONPY	Q8ral2 pongo pygma
5	2077	93.1	656	2 Q8R223_MOUSE	Q8r223 mus musculus
6	812.5	36.4	711	2 Q5GLX3_CIOIN	Q5glx3 ciona intes
7	811.5	36.4	492	2 Q5EB10_MOUSE	Q5eb10 mus musculus
8	805	36.1	735	2 Q5YLV5_XENLA	Q5ylv5 xenopus lae
9	800	35.8	585	2 Q6GQ18_XENLA	Q6gq18 xenopus lae
10	800	35.8	720	2 Q4V7Y0_XENLA	Q4v7y0 xenopus lae
11	800	35.8	754	2 Q8AYP0_ANGJA	Q8ayp0 anguilla ja
12	797.5	35.7	713	2 Q6L968_ANGJA	Q6l968 anguilla ja
13	789	35.3	766	2 Q6XDT1_CIOIN	Q6xdt1 ciona intes
14	785	35.2	739	2 Q7T2C4_BRARE	Q7t2c4 brachydanio
15	781	35.0	726	2 Q6DJN0_XENLA	Q6djn0 xenopus lae
16	774	34.7	739	1 S26A2_MOUSE	Q62273 mus musculus
17	774	34.7	739	2 Q8R213_MOUSE	Q8r213 mus musculus
18	774	34.7	739	2 Q543D6_MOUSE	Q543d6 mus musculus
19	772	34.6	714	2 Q7T2N6_BRARE	Q7t2n6 brachydanio
20	765.5	34.3	804	2 Q6L967_ANGJA	Q6l967 anguilla ja
21	763.5	34.2	563	2 Q8NGQ4_HUMAN	Q8ngq4 homo sapien
22	762.5	34.2	757	2 Q8BGH1_RABIT	Q8bgh1 oryctolagus
23	759	34.0	639	1 S26A2_RAT	Q70531 rattus norv
24	757	33.9	729	2 Q4SVG1_TETNG	Q4svg1 tetraodon n
25	755.5	33.8	734	2 Q3BEG8_BOVIN	Q3beg8 bos taurus
26	753.5	33.8	739	1 S26A2_HUMAN	P50443 homo sapien
27	750.5	33.6	736	2 Q6SAC2_HORSE	Q6sac2 equus cabal
28	748.5	33.5	595	2 Q8RGV3_BRARE	Q8rgv3 brachydanio
29	747.5	33.5	780	1 P6ND_HUMAN	Q43511 homo sapien
30	743.5	33.3	733	2 Q69DJ1_BUBBU	Q69dj1 bubalus bub
31	743.5	33.3	780	2 Q5NVA6_PONPY	Q5nva6 pongo pygma

32 742.5 33.3 713 2 Q5U500\_XENLA Q5u500 xenopus lae  
33 742.5 33.3 719 2 Q6XQ07\_XENLA Q6xq07 xenopus lae  
34 741.5 33.2 764 1 S26A3\_HUMAN S26a3 homo sapien  
35 739 33.1 651 2 Q6MZW7\_HUMAN Q6mzw7 homo sapien  
36 738 33.1 757 1 S26A3\_MOUSE S26a3 mus musculus  
37 735.5 33.0 734 2 Q8GJY3\_SHEEP Q8gjj3 ovis aries  
38 735 32.9 727 2 Q5R7H3\_PONPY Q5r7h3 pongo pygma  
39 734 32.9 753 2 Q5GM09\_PIG Q5gm09 sus scrofa  
40 732 32.8 738 2 Q9NQ11\_HUMAN Q9nq11 homo sapien  
41 732 32.8 758 2 Q548A7\_HUMAN Q548a7 homo sapien  
42 732 32.8 759 1 S26A6\_HUMAN S26a6 homo sapien  
43 726.5 32.5 735 2 Q812E2\_MOUSE Q812e2 mus musculus  
44 726 32.5 757 1 S26A3\_RAT Q924c9 rattus norv  
45 725.5 32.5 735 2 Q8CJD0\_MOUSE Q8cjd0 mus musculus

#### ALIGNMENTS

RESULT 1  
Q96RN2\_HUMAN  
ID Q96RN2\_HUMAN PRELIMINARY; PRT; 656 AA.  
AC Q96RN2\_HUMAN  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DB Anion transporter (SLC26A7 protein).  
GN Name=SLC26A7; Synonyms=SUT2/SLC26A7;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=21950770; PubMed=11834742; DOI=10.1074/jbc.M11802200;  
RA Lohi H., Kujala M., Makela S., Lehtonen E., Kestila M.,  
Saarialho-Kere U., Markovich D., Kere J.;  
RT "Functional characterization of three novel tissue-specific anion  
exchangers SLC26A7, -A8, and -A9".  
RL J.-Biol.-Chem.--277.14246-14254(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX TISSUE=Kidney;  
RA MEDLINE=21818883; PubMed=11829495; DOI=10.1006/geno.2002.6689;  
RA Vincourt J.B., Jullien D., Kossida S., Amalric P., Girard J.P.;  
RT "Molecular cloning of SLC26A7, a novel member of the SLC26  
sulfate/anion transporter family, from high endothelial venules and  
kidney".  
RL Genomics-79:249-256(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX TISSUE=Kidney;  
RA Girard J.;  
RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins L., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalilov D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Kidney;  
RC NIH MGC Project;  
RG TISSUE=Placenta;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF331521; AK95665.1; -; mRNA.  
DR EMBL; AJ413228; CAC88370.1; -; mRNA.  
DR EMBL; BC094730; AAH94730.1; -; mRNA.  
DR EMBL; ENSG00000147606; Homo sapiens.  
DR Ensembl; ENSG00000147606; Homo sapiens.  
DR InterPro; IPR002645; STAS.  
DR InterPro; IPR011547; Sulph\_transp.  
DR Pfam; PF01740; STAS; 1.  
DR Pfam; PF00916; Sulfate\_transp; 1.  
DR PROSITE; PS50801; STAS; 1.  
SQ SEQUENCE 656 AA; 72213 MW; 9P8D2E140B5C00CB CRC64;  
  
Query Match 100.0%; Score 2232; DB 2; Length 656;  
Best Local Similarity 100.0%; Pred. No. 4.5e-156;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MTGAKKKKSLMSKMTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60  
Db 1 MTGAKKKKSLMSKMTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60  
  
Qy 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLSANAVRIVP 120  
Db 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLSANAVRIVP 120  
  
Qy 121 QNMQLTTQSTNSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVTEPVIS 180  
Db 121 QNMQLTTQSTNSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVTEPVIS 180  
  
Qy 181 AMTTGAATHVTSQVKYLLGKMPYISGPIGFFIYAVFENIKSVRLAALLLSLSTV 240  
Db 181 AMTTGAATHVTSQVKYLLGKMPYISGPIGFFIYAVFENIKSVRLAALLLSLSTV 240  
  
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGISPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGISPRAP 300  
  
Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISDDNQEFHLAHLGSLNIVSFPFCI 360  
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISDDNQEFHLAHLGSLNIVSFPFCI 360  
  
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIVIIYAIIGPLLYWLPVCVLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIVIIYAIIGPLLYWLPVCVLASIIIVVGLKGM 420  
  
Qy 421 LIQFRDLKKYNNVDKIDWG 439  
Db 421 LIQFRDLKKYNNVDKIDWG 439

## RESULT 2

Q8TE53 HUMAN  
ID Q8TE53 HUMAN PRELIMINARY; PRT; 663 AA.  
AC Q8TE53;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Anion transporter.  
GN Name=SLC26A7; Synonyms=SLC26A7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;

RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Kidney;  
RC MBLINE=21818883; PubMed=11829495; DOI=10.1006/geno.2002.6689;  
RX Vincourt J.B., Jullien D., Kossida S., Amalric F., Girard J.P.;  
RT "Molecular cloning of SLC26A7, a novel member of the SLC26  
RT sulfate/anion transporter family, from high endothelial venules and  
RT kidney.",  
RL Genomics 79:249-256 (2002).  
RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Kidney;  
RC TISSUE=Kidney;  
RA Vincourt J., Jullien D.L., Amalric F., Girard J.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ413230; CAC88372.1; -; mRNA.  
DR Ensembl; ENSG00000147606; Homo sapiens.  
DR HGNC; HGNC:14467; SLC26A7.  
DR GO; GO:0016021; C:integral  
DR InterPro; IPR011547; Sulph\_transp.  
DR Pfam; PF00916; Sulfate\_transp; 1.  
DR PROSITE; PS50801; STAS; 1.  
SQ SEQUENCE 663 AA; 73032 MW; 1F8967404F54DC9F CRC64;  
  
Query Match 100.0%; Score 2232; DB 2; Length 663;  
Best Local Similarity 100.0%; Pred. No. 4.5e-156;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MTGAKKKKSLMSKMTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60  
Db 1 MTGAKKKKSLMSKMTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60  
  
Qy 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLSANAVRIVP 120  
Db 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLSANAVRIVP 120  
  
Qy 121 QNMQLTTQSTNSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVTEPVIS 180  
Db 121 QNMQLTTQSTNSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVTEPVIS 180  
  
Qy 181 AMTTGAATHVTSQVKYLLGKMPYISGPIGFFIYAVFENIKSVRLAALLLSLSTV 240  
Db 181 AMTTGAATHVTSQVKYLLGKMPYISGPIGFFIYAVFENIKSVRLAALLLSLSTV 240  
  
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGISPRAP 300  
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGISPRAP 300  
  
Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISDDNQEFHLAHLGSLNIVSFPFCI 360  
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISDDNQEFHLAHLGSLNIVSFPFCI 360  
  
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIVIIYAIIGPLLYWLPVCVLASIIIVVGLKGM 420  
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIVIIYAIIGPLLYWLPVCVLASIIIVVGLKGM 420  
  
Qy 421 LIQFRDLKKYNNVDKIDWG 439  
Db 421 LIQFRDLKKYNNVDKIDWG 439

## RESULT 3

Q8TE54 HUMAN  
ID Q8TE54 HUMAN PRELIMINARY; PRT; 656 AA.  
AC Q8TE54;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Anion transporter.  
GN Name=SLC26A7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:40:15 ; Search time 188 Seconds  
(without alignments)  
1025.996 Million cell updates/sec

Title: US-10-761-033-2\_COPY\_1\_439

Perfect score: 2232  
Sequence: 1 MTGAKRKXKXKMLNSKMTTPQ.....MLIQFRDLKXVNVYKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21:.\*  
1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*  
9: geneseqp2005s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2232	100.0	598	AAU74917	Aau74917 Novel hum
2	2232	100.0	605	AAU74915	Aau74915 Novel hum
3	2232	100.0	621	AAU74913	Aau74913 Novel hum
4	2232	100.0	656	AAU74916	Aau74916 Novel hum
5	2232	100.0	663	AAU74914	Aau74914 Novel hum
6	2232	100.0	663	ADC35057	Adc35057 Human SUT
7	2232	100.0	679	AAU74912	Aau74912 Novel hum
8	2227	99.8	656	AAE22123	Aae22123 Human ani
9	2227	99.8	656	ADC35056	Adc35056 Human SUT
10	2077	93.1	656	ABR84485	AbR84485 Human ani
11	2077	93.1	656	ABR84486	AbR84486 Mouse ani
12	2065.5	92.5	671	ABM82599	Abm82599 Human dia
13	2064.5	92.5	671	ABG61534	Abg61534 Human tra
14	938.5	42.0	431	ABG06299	Abg06299 Novel hum
15	825.5	37.0	555	ABB99660	Abb99660 Amino aci
16	805	36.1	735	ADM56357	Adm56357 Frog SLC2
17	774	34.7	739	ADM56280	Adm56280 Mouse SLC
18	759	34.0	739	AAV08929	Aav08929 Rat DTPST
19	759	34.0	739	ADD46847	Add46847 Rat Prote
20	753.5	33.8	739	AAV08928	Aav08928 Human DTD
21	753.5	33.8	739	AAU74765	Aau74765 Human SLC
22	753.5	33.8	739	AAU74763	Aau74763 Human SLC
23	753.5	33.8	739	AAU74762	Aau74762 Human SLC
24	753.5	33.8	739	AAU74764	Aau74764 Human SLC

ALIGNMENTS

RESULT 1  
AAU74917  
ID AAU74917 standard; protein; 598 AA.  
XX  
AC AAU74917;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Novel human protein (NHP) sequence #6.  
XX  
KW Novel human protein; NHP; transporter protein; polymorphism;  
KW mental disorder; biological disease; medical disorder.  
XX  
OS Homo sapiens.  
XX  
FN WO200194583-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 06-JUN-2001; 2001WO-US018393.  
XX  
PR 07-JUN-2000; 2000US-0210045P.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Walke DW, Scoville J;  
XX  
DR WPI; 2002-147673/19.  
DR N-PSDB; ABK12981.  
XX  
PT New human polynucleotides encoding proteins that share sequence  
PT similarity with mammalian transporter proteins, useful for determining  
PT genomic structures, identifying polymorphisms, or as reagents for  
PT diagnosis or drug screening.  
XX  
Claim 1; Page 46-47; 48pp; English.  
XX  
The present invention relates to new isolated nucleic acid molecules  
XX comprising a nucleotide sequence encoding a protein that shares sequence  
XX similarity with mammalian transporter proteins. The invention also  
XX relates to a nucleotide sequence that hybridises under stringent  
XX conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully  
XX defined in the specification or its complement. The isolated nucleic acid  
XX and the protein it encodes are useful for identifying a coding sequence  
XX and mapping a unique gene to a particular chromosome, identifying and  
XX characterising the temporal and tissue specific expression of a gene,  
XX screening a human genomic library, determining the genomic structure of a

25 753.5 33.8 739 6 AAB36029 Aae36029 Human dia  
26 753.5 33.8 739 7 ADB75545 Adb75545 Prostate  
27 753.5 33.8 739 7 ADD46849 Add46849 Human Pro  
28 753.5 33.8 739 7 ADM56364 Adm56364 Human SLC  
29 753.5 33.8 739 8 ADP43501 Adp43501 Human SLC  
30 753.5 33.8 739 9 AEA04531 Aea04531 Human pro  
31 747.5 33.5 780 4 AAM51696 Aam51696 Human pen  
32 747.5 33.5 780 7 ADN95180 Adn95180 Human BEC  
33 747.5 33.5 780 8 ADJ74751 Adj74751 Human ge  
34 747.5 33.5 780 8 ADJ75415 Adj75415 Marker ge  
35 747.5 33.5 790 5 ABG61914 Abg61914 Prostate  
36 747.5 33.5 790 7 ADN39612 Adn39612 Cancer/an  
37 744 33.3 466 8 ADH22527 Adh22527 Human tra  
38 742.5 33.3 719 7 ADM56355 Adm56355 Frog SLC2  
39 741.5 33.2 686 8 ADH22525 Adh22525 Human tra  
40 741.5 33.2 686 8 ABM84848 Abm84848 Human dia  
41 741.5 33.2 764 2 AAR60568 Aar60568 Down-regu  
42 741.5 33.2 764 6 ADA10955 Ada10955 Human CDN  
43 741.5 33.2 764 8 ADG47981 Adg47981 Human col  
44 734 32.9 753 7 ADM56359 Adm56359 Pig SLC26  
45 732 32.8 714 8 ADU48435 Adu48435 Human put

CC given locus/allele or designing diagnostic tests. The nucleic acids and  
CC proteins are particularly useful for identifying polymorphisms and in  
CC amplification assays to detect mutations within the exons, introns and  
CC splice sites that can be used in diagnostics and pharmacogenomics. These  
CC are also useful for generating antibodies, as reagents in diagnostic  
CC assays or for identifying other cellular gene products related to novel  
CC human proteins. The nucleotides can be used as reagents in assays for  
CC screening for compounds that can be employed as pharmaceutical reagents  
CC useful in the therapeutic treatment of mental, biological or medical  
CC disorders and diseases. The present amino acid sequence represents novel  
CC human protein #6 that is one of several (AAU74912- AAU74917) novel human  
CC proteins (NHP) of the invention  
xx  
SQ Sequence 598 AA;

07-JUN-2000; 2000US-0210045P.  
(LEXI-) LEXICON GENETICS INC.  
Walke DW, Scoville J;  
WPI; 2002-147673/19.  
N-PSDB; ABK12979.  
New human polynucleotides encoding proteins that share sequence  
similarity with mammalian transporter proteins, useful for determining  
genomic structures, identifying polymorphisms, or as reagents for  
diagnosis or drug screening.  
Claim 1; Page 42-43; 48pp; English.  
The present invention relates to new isolated nucleic acid molecules  
comprising a nucleotide sequence encoding a protein that shares sequence  
similarity with mammalian transporter proteins. The invention also  
relates to a nucleotide sequence that hybridizes under stringent  
conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully  
defined in the specification or its complement. The isolated nucleic acid  
and the protein it encodes are useful for identifying a coding sequence  
and mapping a unique gene to a particular chromosome, identifying and  
characterising the temporal and tissue specific expression of a gene,  
screening a human genomic library, determining the genomic structure of a  
given locus/allele or designing diagnostic tests. The nucleic acids and  
proteins are particularly useful for identifying polymorphisms and in  
amplification assays to detect mutations within the exons, introns and  
splice sites that can be used in diagnostics and pharmacogenomics. These  
are also useful for generating antibodies, as reagents in diagnostic  
assays or for identifying other cellular gene products related to novel  
human proteins. The nucleotides can be used as reagents in assays for  
screening for compounds that can be employed as pharmaceutical reagents  
useful in the therapeutic treatment of mental, biological or medical  
disorders and diseases. The present amino acid sequence represents novel  
human protein #4 that is one of several (AAU74912- AAU74917) novel human  
proteins (NHP) of the invention  
Sequence 605 AA;  
SQ

301	PMNLSAVITEAFGVALGVYVNSLAQGSACKRISIDNQDEFLAHGUSNIVSSFFFCI	360
361	PSAAMGRTRAGLYSTGAKTQVACLISCIFVLVIYATGPLLXYLPMCVTLASIIIVGLKGM	420
361	PSAAMGRTRAGLYSTGAKTQVACLISCIFVLVIYATGPLLXYLPMCVTLASIIIVGLKGM	420
421	LIQFRDLKKYMNVDKIDWG	439
421	LIQFRDLKKYMNVDKIDWG	439